SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(116th Edition)

AUGUST 2022

BY: CENTRAL LIBRARY

DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY *DELHI COLLEGE OF ENGINEERING)*
GOVT. OF N.C.T. OF DELHI
SHAHBAD DAULATPUR, MAIN BAWANA ROAD
DELHI 110042

PREFACE

This is the **One Hundred Sixteenth** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **August, 2022**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A CONJECTURE ON H3(1) FOR CERTAIN STARLIKE FUNCTIONS, ***8.1.NEHA VERMA*** AND ***3.S. SIVAPRASAD KUMAR***, Applied Mathematics, DTU

2. A Hybrid Finite Difference Method for Singularly Perturbed Delay Partial Differential Equations with Discontinuous Coeffifficient and Source, ***8.Nitika Sharma*** and ***3.Aditya Kaushik***, Applied Mathematics, DTU

3. A Performance Perspective of Live Migration of Virtual Machine in Cloud Data Center with Future Directions, Divya Kapil and ***7.1.Saurabh Kumar Mishra***, CSE, DTU

4. A picture of Delhi's regional air quality during diminished anthropogenic activities in the COVID-19 era, ***6.Vignesh Mohan*** and ***3.Rajeev Kumar Mishra***, Envt., DTU

5. A player unknown's battlegrounds ranking based optimization technique for power system optimization problem, Kapil Deo Bodha, V. Mukherjee and ***3.Vinod Kumar Yadav***, Electrical, DTU

6. A reformulation of collision avoidance algorithm based on artificial potential fields for fixed-wing UAVs in a dynamic environment, ***8.Astik Srivastava*** and PB Sujit, Applied Physics, DTU

7. A Review on Opinion Leader Detection and its Applications, ***6.Sonia***, ***3.Kapil Sharma*** and ***6.Monika Bajaj***, IT and CSE, DTU

8. A review on self-healing polymers for applications in spacecraft and construction of roads, Megha Goyal, ***7.1.Satya Narayan Agarwal*** and Nitu Bhatnagar, ECE, DTU

9. Algal Bioremediation of heavy metals: an insight into removal mechanisms, recovery of by-products, challenges, and future opportunities, ***6.Mohita Chugh***, ***7.Lakhan Kumar***, Maulin P Shah and ***3.Navneeta Bharadvaja***, Biotechnology, DTU

10. An alumni-based collaborative model to strengthen academia and industry partnership: The current challenges and strengths, ***3.Ruchika Malhotra***, ***6.Massoud Massoudi*** and ***3.Rajni Jindal***, CSE, DTU

11. Analysis of municipal solid waste management strategies in the city of Greater Noida, V. Varna Vishakar, B. Dhilipkumar, Ayush Jain and ***7.1.Zohaib Ahmed Khan***, Civil, DTU

12. Analysis of organic rankine cycle powered by N-number of solar collectors in series, Deepak Tiwari, Deepali Atheaya, ***3.Anil Kumar*** and Nishant Kumar, Mechanical and Environment, DTU

13. Being ethically resilient during COVID-19: a cross-sectional study of Indian supply chain companies, ***6.Shivangi Verma***, ***3.Naval Garg*** and Thangaraja Arumugam, USME, DTU

14. Circularity issues and blockchain technology in the auto industry, ***6.Syed Wasiul Hasan Rizvi***, ***3.Saurabh Agrawal*** and ***3.Qasim Murtaza***, Mechanical and DSM, DTU

15. COEFFICIENT PROBLEMS FOR CERTAIN CLOSE-TO-CONVEX FUNCTIONS, ***6.MRIDULA MUNDALIA*** AND ***3.S. SIVAPRASAD KUMAR***, Applied Mathematics, DTU

16. Comparative Analysis of Different hand gestures Detection and Recognition Techniques, ***8.Rajat Kumar***, ***8.Puneet Sharma*** and ***3.Dr. Bindu Verma***, IT, DTU

17. Deep learning based Novel Vehicle Detection and Counting System, ***3.Varsha Sisaudia*** and ***6.Sachin Kumar Giri***, IT, DTU

18. DESIGN AND DEVELOPMENT OF AN AUTONOMOUS UNDERWATER VEHICLE - LAPRAS 2.0, ***8.Sriram Chava***, ***8.Saurav Ranjan***, ***8.Shivang Dudani***, ***8.Utkarsh Verma***, ***8.Shorya Kumar***, ***8.Kritika Bansal***, Rajat, Sundeep Chand, S. Suchit, Chirag Garg, Suryansh Gangwar, Ishpreet Singh Luthra, Shivam Rastogi, Abhinav Agrawal, Rakshit Garg, Dhruv Upreti, Aaryaveer Bhardwaj, Aaryan Mendiratta, Dev Verma, Harjas Singh, ***8.Vidushi Saxena***, ***8.Hriday Bhardwaj***, ***8.Parth Sharma***, ***8.Anupam Maurya***, ***8.Shubh Mittal***, Electrical and Electronics, DTU

19. Disease Detection in Maize Plant using Deep Convolutional Neural Network, ***8.Sohum Gupta***, ***8.Josh Vishnoi*** and ***3.A.S. Rao***, Applied Physics, DTU

20. Diurnal and Seasonal Variation of Atmospheric Particulate Matter and Trace Gases in Industrial Area of Delhi: A Study, Papiya Mandal, ***3.Raju Sarkar***, Neel Kamal, Manob Das and ***6.Anubha Mandal***, Civil, DTU

21. DNS-Tunnet: A Hybrid Approach for DNS Tunneling Detection, ***8.Anushka Lal***, ***8.Ankit Prasad***, ***8.Ankit Kumar*** and ***3.Shailender Kumar***, CSE, DTU

22. Domain-Specific Hybrid BERT based System for Automatic Short Answer Grading, ***8.Jai Garg***, ***8.Jatin Papreja***, ***8.Kumar Apurva*** and ***3.Dr. Goonjan Jain***, Applied Mathematics, DTU

23. Dual-Band Biosensing with van der Waals Assisted Optical Metasurfaces, Rashmi Kumari, ***7.1.Shubhanshi Sharma***, Shailendra Kumar Varshney and Basudev Lahiri, ECE, DTU

24. Ecosystem Based Adaptation and Disaster Risk Reduction in Indian Himalayan Region of Darjeeling, ***3.Raju Sarkar***, ***8.Saif Alam***, ***8.Sanchit Jain*** and ***8.Sarthak Saharan***, Civil, DTU

25. Effects of Different Laser Shape Pulses on Laser Wakefield Acceleration, ***8.Vikrant K Patel***, ***8.Raunaq Singh***, ***8.Nipul Agarwal***, ***6.Monika Yadav***, ***3.Suresh C Sharma***, Applied Physics, DTU

26. Energy input–output analysis and greenhouse gas emission in okra and tomato production in Chotanagpur plateau region of India, B. Sarkar, B. Das, P. K. Sundaram, S. S. Mali, A. P. Anurag, A. Upadhyaya, N. Chandra, B. P. Bhatt and ***3.A. Kumar***, Environment, DTU

27. Exploring dimensions of the customer experience at budget hotels during the COVID-19 pandemic: a netnography approach, Smitha Girija, Devika Rani Sharma and ***6.Vaishali Kaushal***, DSM, DTU

28. Feature Ranking using Statistical Techniques for Computer Networks Intrusion Detection, ***8.Yash Sharma***, ***8.Somya Sharma*** and ***3.Anshul Arora***, Applied Mathematics, DTU

29. Graphene and graphene nanohybrid composites-based electrodes for physiological sensing applications, ***6.Bani Gandhi*** and ***3.Nallanthighal Srinivasa Raghava***, Electronics, DTU

30. Grey wolf and Fuzzy logic based enhanced clustering approach for energy efficient operation of wireless sensor networks, ***8.Hardik Rohilla***, ***8.Himanshu***, ***8.Harish Kumar*** and ***3.Pawan Singh Mehra***, CSE, DTU

31. Groundwater quality assessment and modeling utilizing water quality index and GIS in Kabul Basin, Afghanistan, ***3.S. K. Singh*** and ***6.Ali Reza Noori***, Envt., DTU

32. Implementation of In-loop Filter based DSC-FLL Control Approach for Power Quality Improvement in Isolated Microgrid, ***6.Sombir***, ***3.Madhusudan Singh*** and Ashutosh K. Giri, Electrical, DTU

33. Opinion Leader Detection in Asian Social Networks using Modified Spider Monkey Optimization, ***3.SANJAY KUMAR***, AKSHI KUMAR, ***8.ABHISHEK MALLIK*** and SAKSHI DHALL, CSE, DTU

34. Optimising IOT Based Smart Home Systems Using Machine-Learning Algorithm, ***8.Aman Kumar***, ***8.Aniket Kumar Biswas***, ***8.Abhishek Kumar*** and ***3.Rajesh Kumar Yadav***, CSE, DTU

35. OTRA Based Voltage Mode Third Order Quadrature Oscillator, ***3.Rajeshwari Pandey***, ***3.Neeta Pandey***, ***8.Gurumurthy Komanapalli*** and Rashika Anurag, Electronics, DTU

36. Parameter Optimization for Enhanced Biodiesel Yield from Linum usitatissimum Oil Through Solar Energy Assistance, ***6.Neeraj Budhraja***, Mechanical, DTU

37. Performance evaluation of an industrial solar dryer in Indian scenario: a techno-economic and environmental analysis, Ankit Srivastava, Abhishek Anand, Amritanshu Shukla, ***3.Anil Kumar*** and Atul Sharma, Mechanical, DTU

38. Performance Evaluation of Four Meta-Heuristic MPPT Algorithms for a DC-DC Converter based PV System under Dynamic Weather Conditions (DWC), ***3.Vinay Kumar***, ***8.Yash Dagar*** and ***3.Prakash Chittora***, Electrical, DTU

39. Phycoremediation of Dyes: An Overview on Mechanism, Challenges, and Prospects, ***6.M. Chugh*** and ***3.N. Bharadvaja***, Biotechnology, DTU

40. Quality of Services in Hybrid SDN (hSDN): A Review, ***6.Samiullah Mehraban*** and ***3.R K Yadav***, CSE, DTU

41. Radiation-Sensitive AlGaN/GaN MOS-HEMT-Based Dosimeter, Ruby Mann, ***3.Sonam Rewari***, Praveen Pal, Shobha Sharma and R. S. Gupta, Electronics, DTU

42. RADIUS CONSTANTS OF SIGMOID STARLIKE FUNCTIONS, ***6.PRIYANKA GOEL*** and ***3.S. SIVAPRASAD KUMAR***, Applied Mathematics,, DTU

43. Radius Properties of Harmonic Mappings with Fixed Analytic Part, ***6.Kamaljeet Gangania***, Applied Mathematics, DTU

44. Real-Time Object Size Dimensioning in Raspberry Pi, ***8.Sahil Khadane***, ***8.Sahil Saini***, ***8.Samarth Mittal*** and ***3.Prof. Kapil Sharma***, IT, DTU

45. Regional Languages Hate Speech Detection System Using IndicBERT and Transfer Learning, ***8.Ayush Tanwar***, CSE, DTU

46. Scientific Journal of Silesian University of Technology. Series Transport, ***6.Mohneesh Kumar Sharma***, Mechanical, DTU

47. SentiGames- A Game Theoretic Approach To Sentiment Analysis, ***3.Minni Jain***, ***8.Banavath Gayathri***, ***8.Ritwik Ranjan*** and ***8.Rohit***, CSE, DTU

48. SIRA: a model for propagation and rumor control with epidemic spreading and immunization for healthcare 5.0, Akshi Kumar, ***8.Nipun Aggarwal*** and ***3.Sanjay Kumar***, CSE, DTU

49. STARLIKE FUNCTIONS ASSOCIATED WITH A NON-MA-MINDA FUNCTION, ***3.S. SIVAPRASAD KUMAR*** AND ***6.POOJA YADAV***, Applied Mathematics, DTU

50. STARLIKE FUNCTIONS ASSOCIATED WITH A PETAL SHAPED DOMAIN, ***6.Kush Arora*** and ***3.S. Sivaprasad Kumar***, Applied Mathematics, DTU

51. STUDY OF BYPRODUCTS FORMED BY PYROLYSIS OF BIOMASS, ***8.Parth Aggarwal*** and ***8.Bhavya Choudhary***, Mechanical, DTU

52. Synthesis of diamagnetic ZnO nano-crystallites via sol-gel method and their photocatalytic activity, ***6.Deepak Kumar*** and ***3.Mohan Singh Mehata***, Applied Physics, DTU

53. The compressive study of energy security prospects in India through Solar Power, Om Prakash, Asim Ahmad, ***3.Anil Kumar***, Rajeshwari Chatterjee, Shubham Sharma, Reza Alayi and Hossein Monfared, Mechanical, DTU

54. THEORY OF CERTAIN NON-UNIVALENT ANALYTIC FUNCTIONS, ***6.KAMALJEET GANGANIA***, Applied Mathematics, DTU

55. Therapeutic applications of mesenchymal cells derived from human placenta A comprehensive meta-review, ***8.Ritika Luthra***, ***8.Saanya Yadav*** and ***3.Smita Rastogi Verma***, Biotechnology, DTU

56. Thermodynamics, kinetics and isotherm studies on the removal of anionic Azo-dye (Congo red) using synthesized Chitosan/Moringa oleifera gum hydrogel composites, Sakshi Nangia, Deeksha Katyal and ***3.Sudhir G. Warkar***, Applied Chemistry, DTU

57. Tyrosinase-functionalized gold nanoparticle-tailored ultrasensitive nanosensing probe for hazardous and nutritional phenolic compounds, Jasmeen Dhiman, Kalyan Vaid, Treesa Johns, Ruchika Maurya, Mahima Arora, ***7.1.Ankita Negi***, Ritika Gupta, Mrinmoy Misra, Ki-Hyun Kim and Vanish Kumar, Biotechnology, DTU

58. Voltage Regulation in a Radial Microgrid with High RES Penetration Approach-Optimum DVR Control, ***8.1.Ritika Gour*** and ***3.Vishal Verma***, Electrical, DTU

59. Wind-Induced Interference Effects on Low-Rise Pitched Roof Structure Based on CFD Simulations, ***8.Amartya Sinha***, ***8.Aditya Kumar Jha*** and ***3.Ritu Raj***, Civil, DTU

60. Workplace Spirituality and Psychological Capital (PsyCap) of the Employees: A Cross-sectional Study, Ayatakshee Sarkar, ***3.Naval Garg*** and Sasmita Palo, USME, DTU

1. *Vice Chancellor*
2. *Pro Vice Chancellor*
3. *Faculty*
4. *Teaching-cum-Research Fellow*
5. *Asst. Librarian*
6. *Research Scholar*
7. *PG Scholar*
8. *Undergraduate Student*

1.1. *Ex Vice chancellor*
2.1. *Ex Pro Vice Chancellor*
3.1. *Ex Faculty*
4.1. *Alumni*
5.1 *Others*
6.1. *Ex Research Scholar*
7.1. *Ex PG Scholar*
8.1. *Ex Undergraduate Student*

arXiv:2208.02975v1 [math.CV] 5 Aug 2022

A CONJECTURE ON $H_3(1)$ FOR CERTAIN STARLIKE FUNCTIONS

NEHA VERMA AND S. SIVAPRASAD KUMAR

Abstract. We prove a conjecture concerning the third Hankel determinant, proposed in "Anal. Math. Phys., https://doi.org/10.1007/s13324-021-00483-7", which states that $|H_3(1)| \leq 1/9$ is sharp for the class $\mathcal{S}^*_\wp = \{zf'(z)/f(z) \prec \varphi(z) := 1 + ze^z\}$. In addition, we also establish bounds for sixth and seventh coefficient, and $|H_4(1)|$ for functions in $\mathcal{S}^*_\wp$. The general bounds for two and three-fold symmetric functions related to the Ma-Minda classes $\mathcal{S}^*(\varphi)$ of starlike functions are also obtained.

1. Introduction

For the given positive integers n and q, the Hankel determinant $H_q(n)$ related to the function $f(z) = z + \sum_{n=2}^{\infty} a_n z^n \in \mathcal{A}$, the class of normalized analytic functions, given by

$$H_q(n) = \begin{vmatrix} a_n & a_{n+1} & \dots & a_{n+q-1} \\ a_{n+1} & a_{n+2} & \dots & a_{n+q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n+q-1} & a_{n+q} & \dots & a_{n+2q-2} \end{vmatrix}$$

where $a_1 = 1$, was defined by Noonan and Thomas [17]. For various choices of q and n, the growth of $H_q(n)$ was explored for many subfamilies of univalent functions. Janteng et al. [5] discovered the sharp estimates on the second Hankel determinant, for the classes of starlike and convex functions. Krishna et al. [9] calculated the best estimates of $H_2(2)$ for the class of Bazilevič functions.

Moreover,

$$H_3(1) := 2a_2a_3a_4 + a_3(a_5 - a_3^2) - a_4^2 - a_2^2a_5 \tag{1.1}$$

is the third order Hankel detrminant. Zaprawa [21] evaluated the non sharp bounds on the third Hankel determinant as $|H_3(1)| \leq 1$ and $|H_3(1)| \leq 49/540$ for the classes of starlike and convex functions, respectively. The presence of higher order coefficients in the expression of $H_3(1)$ makes it difficult to solve, and the sharpness of the solution is crucial. Overcoming the challenges, Kowaczk et al. [7] established the sharp bound $|H_3(1)| \leq 4/135$ for the class of convex functions in 2018. Kwon et al. [12] provided the best known estimate for starlike functions given by $|H_3(1)| \leq 4/135$. Later on, in 2018, Lecko et al. [13] proved that $|H_3(1)| \leq 1/9$ is the sharp bound for starlike functions of order $1/2$. By choosing specific values of $\varphi(z)$ in the class $\mathcal{S}^*(\varphi)$ of Ma-Minda [16] defined by

$$\mathcal{S}^*(\varphi) = \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec \varphi(z) \right\}, \tag{1.2}$$

authors in [1] have obtained the sharp estimates for $|H_3(1)| \leq 1/36$ for the choice $\varphi(z) = \sqrt{1+z}$. Since, a proper, careful and infact precise re-arrangement of terms is highly required to obtain the best possible bound and it results in few research articles, which are available in this area for the sharp bound of $H_3(1)$, see [1, 8, 20]. The most important step in obtaining the sharp bound of $H_3(1)$ was to rewrite equation (1.1) in terms of Caratheódory coefficients, especially p_1, p_2, p_3, p_4 and p_5, where $p_i's$ are coefficients of the class $\mathcal{P} := \{p(z) = 1 + \sum_{n=1}^{\infty} p_n z^n : \operatorname{Re} p(z) > 0\}$. Libera

2010 *Mathematics Subject Classification.* 30C45, 30C50.

Key words and phrases. starlike function, subordination, Hankel.

and Zlotkiewicz presented the formula for p_2, p_3 in [14, 15] and Kwon et al. gave the expression for p_4 in [11], which is stated below in the form of a lemma:

Lemma 1.1. *Let $p(z) = 1 + \sum_{n=1}^{\infty} a_n z^n \in \mathcal{P}$. Then,*

$$2p_2 = p_1^2 + \gamma(4 - p_1^2), \tag{1.3}$$

$$4p_3 = p_1^3 + 2p_1(4 - p_1^2)\gamma - p_1(4 - p_1^2)\gamma^2 + 2(4 - p_1^2)(1 - |\gamma|^2)\eta \tag{1.4}$$

and

$$8p_4 = p_1^4 + (4-p_1^2)\gamma(p_1^2(\gamma^2 - 3\gamma + 3) + 4\gamma) - 4(4-p_1^2)(1-|\gamma|^2)(p_1(\gamma-1)\eta + \bar{\gamma}\eta^2 - (1-|\eta|^2)\rho), \tag{1.5}$$

for some γ, η and ρ such that $|\gamma| \leq 1$, $|\eta| \leq 1$ and $|\rho| \leq 1$.

In 2021, Kumar and Gangania [10], introduced a new class $\mathcal{S}^*_{\wp}$ by choosing $\varphi(z) = 1 + ze^z$ in (1.2) defined as

$$\mathcal{S}^*_{\wp} = \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec 1 + ze^z \right\}.$$

They used the similar strategy and obtained the bound as, $|H_3(1)| \leq 0.150627$ for $\mathcal{S}^*_{\wp}$ and also proposed a conjecture which is stated as follows:

Conjecture 1.2. [10, Page no. 33] If $f \in \mathcal{S}^*_{\wp}$, then the sharp bound for the third Hankel determinant is given by

$$|H_3(1)| \leq \frac{1}{9} \approx 0.1111\ldots,$$

with the extremal function $f(z) = z\exp\left(\frac{1}{3}(e^{z^3} - 1)\right) = z + \frac{1}{3}z^4 + \frac{2}{9}z^7 + \cdots$.

In this article, together with the proof of conjecture, the estimates for a_6 and a_7 in association with the fourth order Hankel determinant for the functions in the class $\mathcal{S}^*_{\wp}$ are obtained. For $\mathcal{S}^*(\varphi)$, the general third Hankel determinant for second and third fold symmetric functions are also estimated.

2. Proof of the Conjecture 1.2

The initial coefficients for the functions in class $\mathcal{S}^*_{\wp}$ given in [10], are as follows:

$$a_2 = \frac{p_1}{2}, \quad a_3 = \frac{1}{4}\left(p_2 + \frac{p_1^2}{2}\right), \quad a_4 = \frac{1}{6}\left(p_3 + \frac{3}{4}p_1p_2\right) \tag{2.1}$$

$$a_5 = \frac{1}{8}\left(\frac{p_1^4}{48} + \frac{p_2^2}{4} + \frac{2p_1p_3}{3} - \frac{p_1^2p_2}{8} + p_4\right) \tag{2.2}$$

$$a_6 = \frac{1}{4}\left(- \frac{p_1^5}{240} + \frac{19p_1^3p_2}{480} - \frac{7p_1p_2^2}{80} - \frac{p_1^2p_3}{15} + \frac{p_2p_3}{6} + \frac{p_1p_4}{4} + \frac{2p_5}{5}\right) \tag{2.3}$$

and

$$a_7 = \frac{1}{4}\left(\frac{17p_1^6}{11520} - \frac{37p_1^4p_2}{1920} + \frac{29p_1^2p_2^2}{480} - \frac{p_2^3}{32} + \frac{13p_1^3p_3}{360} - \frac{p_1p_2p_3}{6} + \frac{p_3^2}{18} - \frac{p_1^2p_4}{16} + \frac{p_2p_4}{8} + \frac{p_1p_5}{5} + \frac{p_6}{3}\right). \tag{2.4}$$

Now, we proceed by providing a positive response to the conjecture 1.2 in the form of a new theorem which states:

Theorem 2.1. *Let $f \in \mathcal{S}^*_{\wp}$. Then,*

$$|H_3(1)| \leq \frac{1}{9}. \tag{2.5}$$

The result is sharp.

Proof. As the Caratheódory class is rotationally invariant, p_1 lies inside the interval [0,2]. On substituting the expressions of a_i's from equations (2.1) and (2.2) in equation (1.1) assuming $p_1 := p$, we get

$$H_3(1) = \frac{1}{9216}\bigg(3p^6 - 12p^4p_2 + 96p^3p_3 - 192pp_2p_3 - 144p^2p_2^2 + 144p^2p_4 + 72p_2^3 - 256p_3^2 + 288p^2p_4\bigg).$$

On simplification using (1.3)-(1.5), we obtain

$$H_3(1) = \frac{1}{9216}\bigg(\nu_1(p,\gamma) + \nu_2(p,\gamma)\eta + \nu_3(p,\gamma)\eta^2 + \phi(p,\gamma,\eta)\rho\bigg)$$

where $\gamma, \eta, \rho \in \mathbb{D}$,

$$\begin{aligned}
\nu_1(p,\gamma) :&= -4p^6 - 25p^2\gamma^2(4-p^2)^2 - 5p^2\gamma^3(4-p^2)^2 + 2p^2\gamma^4(4-p^2)^2 + 5p^4\gamma(4-p^2)\\
&\quad + 36\gamma^3(4-p^2)^2 - 16p^4\gamma^2(4-p^2),\\
\nu_2(p,\gamma) :&= 8(1-|\gamma|^2)(4-p^2)(4p^3 - (4-p^2)(p\gamma + 10p\gamma^2)),\\
\nu_3(p,\gamma) :&= 8(1-|\gamma|^2)(4-p^2)^2(-8-|\gamma|^2)\\
\phi(p,\gamma,\eta) :&= 72(1-|\gamma|^2)(4-p^2)^2(1-|\eta|^2)\gamma.
\end{aligned}$$

On taking $x = |\gamma|, y = |\eta|$ and using the fact that $|\rho| \leq 1$, we have

$$|H_3(1)| \leq \frac{1}{9216}\bigg(|\nu_1(p,\gamma)| + |\nu_2(p,\gamma)|y + |\nu_3(p,\gamma)|y^2 + |\phi(p,\gamma,\eta)|\bigg) \leq B(p,x,y),$$

where

$$B(p,x,y) = \frac{1}{9216}\bigg(b_1(p,x) + b_2(p,x)y + b_3(p,x)y^2 + b_4(p,x)(1-y^2)\bigg) \tag{2.6}$$

with

$$\begin{aligned}
b_1(p,x) :&= 4p^6 + 25p^2x^2(4-p^2)^2 + 5p^2x^3(4-p^2)^2 + 2p^2x^4(4-p^2)^2 + 5p^4x(4-p^2)\\
&\quad + 36x^3(4-p^2)^2 + 16p^4x^2(4-p^2),\\
b_2(p,x) :&= 8(1-x^2)(4-p^2)(4p^3 + (4-p^2)(px + 10px^2)),\\
b_3(p,x) :&= 8(1-x^2)(4-p^2)^2(8+x^2),\\
b_4(p,x) :&= 72(1-x^2)(4-p^2)^2x.
\end{aligned}$$

We must maximise $B(p,x,y)$ in the cuboid $V : [0,2] \times [0,1] \times [0,1]$. So, we use the maximum values in the interiors of the six faces, the twelve edges, and the interior of V.

(1) To begin with, all the interior points of V are taken into consideration. On partially differentiating equation (2.6) with respect to y to determine its points of maxima in the interior of V by considering $(p,x,y) \in (0,2) \times (0,1) \times (0,1)$. We get

$$\frac{\partial B}{\partial y} = \frac{1}{1152}(4-p^2)(1-x^2)\bigg(4p^3 + (4-p^2)(px(1+10x) - 18xy + 2y(8+x^2))\bigg).$$

Now $\dfrac{\partial B}{\partial y} = 0$ gives

$$y = y_0 := \frac{4px(1+10x) - p^3(10x^2+x-4)}{2(x-1)(x-8)(p^2-4)}.$$

We note that $y_0 \in (0,1)$ assuring the existence of critical points and it is possible when

$$4px(1+10x) - p^3(10x^2+x-4) + 64 - 72x + 2x^2(4-p^2) < 2p^2(8-9x). \tag{2.7}$$

Now we will apply the hit and trial method to find solutions that satisfy the inequalities (2.7) for the existence of critical point. If we assume that p tends to 0 and 2, there is no $x \in (0,1)$ that can satisfy the equation (2.7). Similarly, if x tends to 0 and 1, there is no $p \in (0,2)$ satisfying equation (2.7). As a result, the function B does not have a critical point in $(0,2) \times (0,1) \times (0,1)$.

(2) Now, the interiors of all the faces of V are taken into consideration.
On the face $p = 0, B(p,x,y)$ becomes

$$d_1(x,y) := B(0,x,y) = \frac{9x^3 + 2(1-x^2)((8+x^2)y^2 + 9x(1-y^2))}{144} \tag{2.8}$$

with $x, y \in (0,1)$. We notice that d_1 has no critical points in $(0,1) \times (0,1)$ as

$$\frac{\partial d_1}{\partial y} = \frac{-(1-x)^2(x+1)(x-8)y}{36} \neq 0, \quad x, y \in (0,1). \tag{2.9}$$

On the face $p = 2, B(p,x,y)$ becomes

$$B(2,x,y) := \frac{1}{36}, \quad x, y \in (0,1). \tag{2.10}$$

On the face $x = 0, B(p,x,y)$ becomes

$$d_2(p,y) := B(p,0,y) = \frac{(p^3 + 16y - 4p^2y)^2}{2304} \tag{2.11}$$

with $y \in (0,1)$ and $p \in (0,2)$. To identify the points of maxima, we solve $\partial d_2/\partial p$ and $\partial d_2/\partial y$. On solving $\partial d_2/\partial y = 0$, we obtain

$$y = -\frac{p^3}{4(4-p^2)} (=: y_1). \tag{2.12}$$

Based on simple calculations, we can conclude that such y_1 does not belong to $(0,1)$. As a result, there is no critical point in $(0,2) \times (0,1)$ for d_2.
On the face $x = 1, B(p,x,y)$ becomes

$$d_3(p,y) := B(p,1,y) = \frac{576 + 224p^2 - 136p^4 + 15p^6}{9216}, \quad p \in (0,2). \tag{2.13}$$

During the calculations, $p_0 := p \approx 0.991758$ comes out to be the critical point when $\partial d_3/\partial p = 0$. Moreover, using basic mathematics, d_3 reaches its maximum value of approximately 0.0736789 at p_0.
On the face $y = 0, B(p,x,y)$ becomes

$$\begin{aligned} B(p,x,0) = \frac{1}{9216}\Big(& 4p^4x(23 - 34x - 19x^2 - 4x^3) + p^6(4 - 5x + 9x^2 + 5x^3 + 2x^4) \\ & + 576x(2 - x^2) + 16p^2x(-36 + 25x + 23x^2 + 2x^3)\Big) \\ & =: d_4(p,x). \end{aligned}$$

On computing,

$$\begin{aligned} \frac{\partial d_4}{\partial x} = \frac{1}{9216}\Big(& -4p^4(-23 + 34x + 19x^2 + 4x^3) + p^6(-5 + 18x + 15x^2 + 8x^3) \\ & - 4p^4x(34 + 38x + 12x^2) + 16p^2(-36 + 25x + 23x^2 + 2x^3) \\ & - 1152x^2 - 576(x^2 - 2) + 16p^2x(25 + 46x + 6x^2)\Big) \end{aligned}$$

and

$$\frac{\partial d_4}{\partial p} = \frac{p}{4608}\Big(16x(-36+25x+23x^2+2x^3) - 8p^2x(-23+34x+19x^2+4x^3) + 3p^4(4-5x+9x^2+5x^3+2x^4)\Big).$$

According to a numerical computation, the critical points obtained by solving the system of equations $\partial d_4/\partial x = 0$ and $\partial d_4/\partial p = 0$ in $(0,2)\times(0,1)$ implies

$$B(0.00115734, 0.816497, 0) \leq 0.0680414.$$

On the face $y = 1, B(p,x,y)$ reduces to

$$\begin{aligned} B(p,x,1) = \frac{1}{9216}\Big(&4p^4(16+5x-48x^2-x^3-6x^4) - 8p^5(4-x-14x^2+x^3+10x^4) \\ &+ p^6(4-5x+9x^2+5x^3+2x^4) + 16p^2(-32+53x^2-13x^3+6x^4) \\ &+ 128px(1+10x-x^2-10x^3) + 64(16-14x^2+9x^3-2x^4) \\ &+ 64p^3(2-x-12x^2+x^3+10x^4)\Big) =: d_5(p,x). \end{aligned}$$

We observe that the system of equations $\partial d_5/\partial x = 0$ and $\partial d_5/\partial p = 0$ has no solution in $(0,2)\times(0,1)$.

(3) Now, we determine the maximum values attained on the edges of the cuboid V by $B(p,x,y)$. We get $B(p,0,0) = c_1(p) := 4p^6/9216$ from equation (2.11). It is apparent that $c_1'(p) = 0$ for $p = \beta_0 := 0$ and $p = \beta_1 := 2$ as points of minima and maxima in the interval $[0,2]$, respectively. Maximum value of $c_1(p)$ is ≈ 0.0277778. Hence,

$$B(p,0,0) \leq 0.0277778.$$

With $y = 1$, we get $B(p,0,1) = c_2(p) := (16-4p^2+p^3)2/2304$ from the equation (2.11). Since $c_2'(p) < 0$ in the range $[0,2]$, $p = 0$ acts as its point of maxima. Thus

$$B(p,0,1) \leq \frac{1}{9}, \quad p \in [0,2].$$

Based on computations, $B(0,0,y)$ in equation (2.11) reaches its maximum value at $y = 1$. This leads to

$$B(0,0,y) \leq \frac{1}{9}, \quad y \in [0,1].$$

Since, the equation (2.13) is free from x, we have $B(p,1,1) = B(p,1,0) = c_3(p) := (15p^6 - 136p^4 + 224p^2 + 576)/9216$. Now, $c_3'(p) = 448p - 544p^3 + 90p^5 = 0$ when $p = \beta_2 := 0$ and $p = \beta_3 := 0.991758$ in the interval $[0,2]$ where β_2 and β_3 are the minimum and maximum points, respectively. Hence

$$B(p,1,1) = B(p,1,0) \leq 0.0736789, \quad p \in [0,2].$$

We get $B(0,1,y) = 1/16$ when $p = 0$ is substituted in equation (2.13). The equation (2.10) do not involve any variables namely p, x and y. Thus, on the edges $p = 2, x = 1; p = 2, x = 0; p = 2, y = 0$; and $p = 2, y = 1$, the maximum value of $B(p,x,y)$ is determined by

$$B(2,1,y) = B(2,0,y) = B(2,x,0) = \frac{1}{36}, \quad x,y \in [0,1].$$

From equation (2.11), we obtain $B(0,0,y) = y^2/36$. A simple calculation shows that

$$B(0,0,y) \leq \frac{1}{36}, \quad y \in [0,1].$$

We get $B(0,x,1) = c_4 := (8-7x^2-x^4)/72$ using equation (2.8). We observe that c_4 is a decreasing function in the range $[0,1]$, and so reaches its maximum value at $x=0$, according to a simple calculation. Hence

$$B(0,x,1) \leq \frac{1}{9}, \quad x \in [0,1].$$

$B(0,x,0) = c_5(x) := x(1-x^2)/8$ is obtained using equation (2.8) once again. We get $c_5'(x) = 0$ after further calculations for $x = x_0 := 1/\sqrt{3}$. Also, $c_5(x)$ is an increasing function in $[0,x_0)$ and decreasing in $(x_0,1]$. As a result, maximum value is attained at x_0. Thus

$$B(0,x,0) \leq 0.0481125, \quad x \in [0,1].$$

Thus, the inequality (2.5) holds for all the cases. The sharpness of the result is governed by the function $f : \mathbb{D} \to \mathbb{C}$ defined as

$$f(z) = z\exp\left(\frac{1}{3}(e^{z^3}-1)\right) = z + \frac{z^4}{3} + \frac{2z^7}{9} + \cdots, \tag{2.14}$$

with $f(0) = 0$ and $f'(0) = 1$. For the values of $a_2 = a_3 = a_5 = 0$ and $a_4 = 1/3$ the function specified in equation (2.14) serves as an extremal function for the bounds of $H_3(1)$.

∎

To prove our next result, we must first recall the following lemma:

Lemma 2.2. *Let $p = 1 + \sum_{n=1}^{\infty} p_n z^n \in \mathcal{P}$. Then*

$$|p_n| \leq 2, \quad n \geq 1, \tag{2.15}$$

$$|p_{n+k} - \mu p_n p_k| \leq \begin{cases} 2, & 0 \leq \mu \leq 1; \\ 2|2\mu - 1|, & elsewhere, \end{cases} \tag{2.16}$$

and

$$|p_1^3 - \mu p_3| \leq \begin{cases} 2|\mu - 4|, & \mu \leq 4/3; \\ 2\mu\sqrt{\dfrac{\mu}{\mu-1}}, & 4/3 < \mu. \end{cases} \tag{2.17}$$

Lemma 2.3. *Let $f \in \mathcal{S}_{\wp}^*$. Then $|a_6| \leq 47/60 \approx 0.7833$ and $|a_7| \leq 503/480 \approx 1.0479$.*

Proof. From equation (2.3), we have

$$\begin{aligned} |a_6| &= \left|\frac{1}{4}\left(-\frac{p_1^5}{240} + \frac{19p_1^3p_2}{480} - \frac{7p_1p_2^2}{80} - \frac{p_1^2p_3}{15} + \frac{p_2p_3}{6} + \frac{p_1p_4}{4} + \frac{2p_5}{5}\right)\right| \\ &= \left|\frac{1}{10}\left(p_5 + \frac{5}{8}p_1p_4\right) - \frac{p_1^2}{60}\left(p_3 - \frac{19}{32}p_1p_2\right) + \frac{p_2}{24}\left(p_3 - \frac{21}{40}p_1p_2\right) - \frac{2}{1920}p_1^5\right|. \end{aligned}$$

By using equation(2.15), (2.16) and triangle inequality, we get

$$|a_6| \leq \frac{47}{60}.$$

Similarly, from equation (2.4), we have

$$\begin{aligned} 46080a_7 = {} & 17p_1^6 - 222p_1^4p_2 + 696p_1^2p_2^2 - 360p_2^3 + 416p_1^3p_3 - 1920p_1p_2p_3 + 640p_3^2 - 720p_1^2p_4 \\ & + 1440p_2p_4 + 2304p_1p_5 + 3840p_6 \end{aligned}$$

or

$$46080|a_7| \leq 3840\left|p_6+\frac{3p_2p_4}{8}\right|+416|p_1|^3\left|p_3-\frac{111p_1p_2}{208}\right|+360|p_2|^2\left|-p_2+\frac{29p_1^2}{15}\right| + 2304|p_1|\left|p_5-\frac{5p_2p_3}{6}\right|+|17p_1^6-720p_1^2p_4|+640|p_3|^2.$$

On using equation (2.15) and (2.16), we get

$$3840\left|p_6+\frac{3p_2p_4}{8}\right|\leq 13440, \quad 416|p_1|^3\left|p_3-\frac{111}{208}p_1p_2\right|\leq 6656, \tag{2.18}$$

$$360|p_2|^2\left|p_2-\frac{29p_1^2}{15}\right|\leq 7008, \quad 2304|p_1|\left|p_5-\frac{5p_2p_3}{6}\right|\leq 9216 \tag{2.19}$$

and

$$|17p_1^6-720p1^2p_4|+640|p_3|^2\leq 11968. \tag{2.20}$$

From equations (2.18), (2.19) and (2.20), we get the desired result. ∎

3. Fourth Hankel Determinant and third hankel for n-fold symmetric functions

3.1. **Fourth Hankel estimation for the class $\mathcal{S}^*_\wp$.** For $q=4$ and $n=1$, the expression of the fourth Hankel determinant can be written as

$$H_4(1)=a_7H_3(1)-a_6Q_1+a_5Q_2-a_4Q_3, \tag{3.1}$$

where

$$Q_1=a_3(a_2a_5-a_3a_4)-a_4(a_5-a_2a_4)+a_6(a_3-a_2^2), \tag{3.2}$$

$$Q_2=a_3(a_3a_5-a_4^2)-a_5(a_5-a_2a_4)+a_6(a_4-a_2a_3), \tag{3.3}$$

and

$$Q_3=a_4(a_3a_5-a_4^2)-a_5(a_2a_5-a_3a_4)+a_6(a_4-a_2a_3). \tag{3.4}$$

To compute the fourth Hankel determinant for the function in class $\mathcal{S}^*_\wp$, we substitute the values of a_i from equations (2.1)-(2.4) in equation (3.2). Upon simplification, we have

$$92160Q_1=27p_1^7-408p_1^5p_2+p_1^2p_2(660p_1^2-1224p_2)+p_1^2(336p_2p_3-1152p_5) + p_2(2304p_5-480p_2p_3)+p_4(1440p_1p_2-1920p_3)+392p_1^4p_3.$$

By applying Lemma 2.2, we arrive at

$$\begin{aligned}92160|Q_1| &\leq |p_1|^5|27p_1^2-408p_2|+|p_1|^2|p_2||660p_1^2-1224p_2|+|p_1|^2|336p_2p_3-1152p_5| \\ &\quad +|p_2||2304p_5-480p_2p_3|+|p_4||1440p_1p_2-1920p_3|+392|p_1|^4|p_3| \\ &\leq 84352.\end{aligned}$$

So, we obtain

$$|Q_1|\leq \frac{659}{720}\approx 0.915278. \tag{3.5}$$

In similar way, we have

$$\begin{aligned}737280Q_2 &= 73p_1^8+1440p_1^5p_3+p_1^2p_4(14400p_2-1920p_1^2)+p_1p_2p_3(4096p_1^2-8448p_2) \\ &\quad +p_1^4p_2(-1692p_2-336p_1^2)+p_5(12288p_3-4608p_1^3)-11520p_4^2+720p_2^4 \\ &\quad -4608p_1^2p_3^2.\end{aligned}$$

Using Lemma 2.2, we get

$$\begin{aligned}737280|Q_2| &\leq |p_1|^5|73p_1^3+1440p_3|+|p_1|^2|p_4||14400p_2-1920p_1^2|+|p_1||p_2||p_3||4096p_1^2-8448p_2|\\&\quad+|p_1|^4|p_2||-1692p_2-336p_1^2|+|p_5||12288p_3-4608p_1^3|+11520|p_4|^2+720|p_2|^4\\&\quad+4608|p_1|^2|p_3|^2\\&\leq 759296+98304\sqrt{\frac{2}{5}}.\end{aligned}$$

So, we get

$$|Q_2|\leq\frac{759296+98304\sqrt{\frac{2}{5}}}{737280}\approx 1.11419. \tag{3.6}$$

Again, by rearrangement of terms, we have,

$$\begin{aligned}4423680Q_3 &= 57p_1^9+288p_1^6p_3+p_1^3p_4(25920p_2-5760p_1^2)+p_1^4p_2(16128p_3-144p_1^3)\\&\quad-p_3^2(20480p_3+6144p_1^3)-p_1^3p_2^2(9216p_2+540p_1^2)+p_1p_5(36864p_3-6912p_1^3)\\&\quad-26496p_1^2p_2^2p_3+12528p_1p_2^4+46080p_2p_3p_4-34560p_1p_4^2-27648p_2^2p_5.\end{aligned}$$

From Lemma 2.2, we get,

$$\begin{aligned}4423680|Q_3| &\leq |p_1|^6|57p_1^3+288p_3|+|p_1|^3|p_4||25920p_2-5760p_1^2|+|p_1|^4|p_2||16128p_3-144p_1^3|\\&\quad+|p_3|^2|20480p_3+6144p_1^3|+|p_1|^3|p_2|^2|9216p_2+540p_1^2|+|p_1||p_5||36864p_3-6912p_1^3|\\&\quad+26496|p_1|^2|p_2|^2|p_3|+12528|p_1||p_2|^4+46080|p_2||p_3||p_4|+34560|p_1||p_4|^2\\&\quad+27648|p_2|^2|p_5|\\&\leq 4029952+1376256\sqrt{\frac{21}{37}}+\frac{1179648}{\sqrt{13}}.\end{aligned}$$

So, we get

$$|Q_3|\leq\frac{4029952+1376256\sqrt{\frac{21}{37}}+\frac{1179648}{\sqrt{13}}}{4423680}\approx 1.21934. \tag{3.7}$$

Based on the above computations, we make the following statement on fourth Hankel determinant:

Theorem 3.1. *Let $f\in\mathcal{S}^*_\wp$. Then $|H_4(1)|\leq 2.54589$.*

3.2. **Third Hankel determinant for 2&3 fold symmetric functions for $\mathcal{S}^*(\varphi)$.** In the recent times, it has been observed that finding the sharp estimates of third Hankel determinant for general Ma-Minda class is not feasible till now. But for some classes, sharp estimates have been obtained, for instance, see [1, 8, 20] and now including Theorem 2.1 as well which motivated us to settle the Conjecture 1.2. Further looking at the difficulty of the general class, we restrict ourselves to answer the problem for the n-fold symmetric functions.

Definition 3.2. [3] A function $f\in\mathcal{A}$ is called n-fold symmetric if $f(e^{2\pi i/n}z)=e^{2\pi i/n}f(z)$ which holds for all $z\in\mathbb{D}$ and n is a natural number. We denote the set of n-fold symmetric functions by $\mathcal{A}^{(n)}$.

Let $f\in\mathcal{A}^{(n)}$, then f has power series expansion

$$f(z)=a_1z+a_{n+1}z^{n+1}+a_{2n+1}z^{2n+1}+\cdots.$$

Therefore, for $f\in\mathcal{A}^{(3)}$ and $f\in\mathcal{A}^{(2)}$ respectively, we have

$$H_3(1)=-a_4^2\quad\text{and}\quad H_3(1)=a_3(a_5-a_3^2). \tag{3.8}$$

Now we conclude this paper with the following result:

Theorem 3.3. *Let $f \in \mathcal{S}^*(\varphi)$. Then*

(1) $\widehat{f} \in \mathcal{S}^{*(3)}(\varphi)$ implies that $|H_3(1)| \leq |B_1|^2/9$.

(2) $\widehat{f} \in \mathcal{S}^{*(2)}(\varphi)$ implies that

$$|H_3(1)| \leq \frac{1}{4}|B_1| \times \begin{cases} \frac{1}{6}(B_2 - \frac{9}{8}B_1^2 + B_1^2), & \frac{9}{4}B_1^2 \leq 2(B_2 + B_1^2 - B_1); \\ \frac{1}{6}B_1, & 2(B_2 + B_1^2 - B_1) \leq \frac{9}{4}B_1^2 \leq 2(B_2 + B_1^2 + B_1); \\ \frac{1}{6}(-B_2 + \frac{9}{8}B_1^2 - B_1^2), & 2(B_2 + B_1^2 + B_1) \leq \frac{9}{4}B_1^2. \end{cases}$$

The estimate in (1) *is sharp.*

Proof. Since $f(z) = z + a_2z^2 + a_3z^3 + \cdots \in \mathcal{S}^*(\varphi)$. Let

$$\varphi(z) = 1 + B_1z + B_2z^2 + B_3z^3 + \cdots$$

and

$$p(z) = \frac{zf'(z)}{f(z)} = 1 + b_1z + b_2z^2 + \cdot.$$

This equation shows that

$$(n-1)a_n = \sum_{k=1}^{n-1} b_k a_{n-k} \quad n > 1. \tag{3.9}$$

Since φ is univalent and $p \prec \varphi$, then the function

$$p_1(z) = \frac{1 + \varphi^{-1}(p(z))}{1 - \varphi^{-1}(p(z))} = 1 + c_1z + c_2z^2 + c_3z^3 + \cdots,$$

belongs to the class $\mathcal{P}$. Or equivalently,

$$p(z) = \varphi\left(\frac{p_1(z) - 1}{p_1(z) + 1}\right).$$

Using the last equation, the coefficient b_i can be expressed in terms of c_i and B_i ($i \in \mathbb{N}$). We have

$$b_1 = \frac{1}{2}B_1c_1, \quad b_2 = \frac{1}{4}((B_2 - B_1)c_1^2 + 2B_1c_2)$$

and

$$b_3 = \frac{1}{8}((B_1 - 2B_2 + B_3)c_1^3 + 4(B_2 - B_1)c_1c_2 + 4B_1c_3).$$

Hence, by using the expressions for b_k in equation (3.9), we obtain

$$a_2 = b_1 = \frac{1}{2}B_1c_1, \quad \text{and} \quad a_3 = \frac{1}{8}((B_1^2 - B_1 + B_2)c_1^2 + 2B_1c_2). \tag{3.10}$$

(1) Since $f(z) = z + a_2z^2 + a_3z^3 + \cdots \in \mathcal{S}^*(\varphi)$ if and only if

$$\widehat{f}(z) = (f(z^3))^{1/3} = z + \beta_4z^4 + \cdots \in \mathcal{S}^{*(3)}(\varphi).$$

We have $\beta_4 = b_1/3$. Hence, for $f \in \mathcal{S}^{*(3)}(\varphi)$, from equation (3.8), we obtain

$$H_3(1) = |\beta_4|^2 = \frac{1}{9}|b_1|^2 = \frac{1}{36}|B_1c_1|^2 \tag{3.11}$$

$$\leq \frac{1}{9}|B_1|^2. \tag{3.12}$$

Also, the result is sharp for $f_0(z) = z\exp\int_0^z \frac{\varphi(t)-1}{t}dt$ and its rotations.

(2) Since $f(z) = z + a_2z^2 + a_3z^3 + \cdots \in \mathcal{S}^*(\varphi)$ if and only if

$$\widehat{f}(z) = (f(z^2))^{1/2} = z + \alpha_3 z^3 + \alpha_5 z^5 \cdots \in \mathcal{S}^{*(2)}(\varphi).$$

Upon comparing the coefficients in the following:

$$z^2 + a_2z^4 + a_3z^6 + \cdots = (z + +\alpha_3 z^3 + \alpha_5 z^5 \cdots)^2,$$

we obtain

$$\alpha_3 = \frac{1}{2}a_2 \quad \text{and} \quad \alpha_5 = \frac{1}{2}a_3 - \frac{1}{8}a_2^2. \tag{3.13}$$

If $f \in \mathcal{S}^{*(2)}(\varphi)$, then from equation (3.8), we have

$$H_3(1) = \alpha_3(\alpha_5 - \alpha_3^2) \tag{3.14}$$

Using equation (3.10),

$$\begin{aligned} |H_3(1)| &= \left|\frac{1}{2}a_2\left(\frac{4a_3 - 3a_2^2}{8}\right)\right| = \left|\frac{a_2}{4}\left(a_3 - \frac{3}{4}a_2^2\right)\right| \\ &\leq \frac{1}{4}|a_2|\left|a_3 - \frac{3}{4}a_2^2\right|. \end{aligned}$$

Now, equation (3.10)and Fekete-Szegö bounds [16, Theorem 3] for $\mu = 3/4$, we get the desired result.

∎

Corollary 3.4. *Let $f \in \mathcal{S}^*(\varphi)$ and $H_3(1)$ is given by equation (3.12). Then*

(1) $\widehat{f} \in \mathcal{S}^{*(3)}(1 + ze^z)$ implies that $|H_3(1)| \leq 1/9$.
(2) $\widehat{f} \in \mathcal{S}^{*(3)}((1 + z)/(1 - z))$ implies that $|H_3(1)| \leq 4/9$.

The sharpness of the bounds follows from [10] and [21] respectively.

Corollary 3.5. *Let $f \in \mathcal{S}^*(\varphi)$ and $H_3(1)$ is given by equation (3.14). Then*

(1) $\widehat{f} \in \mathcal{S}^{*(2)}(1 + ze^z)$ implies that $|H_3(1)| \leq 1/24$.
(2) $\widehat{f} \in \mathcal{S}^{*(2)}((1 + z)/(1 - z))$ implies that $|H_3(1)| \leq 1/6$.

Remark 3.6. We observe that the bounds obtained in Corollary 3.5 are close to the sharp values and are still open for their sharpness.

Acknowledgment. Neha is thankful to the Department of Applied Mathematics, Delhi Technological University, New Delhi-110042 for providing Research Fellowship.

References

[1] S. Banga and S. S. Kumar, The sharp bounds of the second and third Hankel determinants for the class $\mathcal{SL}^*$, Math. Slovaca **70** (2020), no. 4, 849–862

[2] S. Fadaei, A. Ebadian and E. A. Adegani, The second Hankel determinant problem for a certain subclass of bi-univalent functions, Afr. Mat. **33** (2022), no. 2, Paper No. 48

[3] A. W Goodman, *Univalent functions*. Vol. I, Mariner Publishing Co., Inc., Tampa, FL, 1983

[4] W. K. Hayman, On the second Hankel determinant of mean univalent functions, Proc. London Math. Soc. (3) **18** (1968), 77–94. MR0219715

[5] A. Janteng, S. A. Halim and M. Darus, Hankel determinant for starlike and convex functions, Int. J. Math. Anal. (Ruse) **1** (2007), no. 13-16, 619–625.

[6] S. Kanas and A. Wisniowska, Conic regions and k-uniform convexity, J. Comput. Appl. Math. **105** (1999), no. 1-2, 327–336

[7] B. Kowalczyk, A. Lecko and Y. J. Sim, The sharp bound for the Hankel determinant of the third kind for convex functions, Bull. Aust. Math. Soc. **97** (2018), no. 3, 435–445.

[8] B. Kowalczyk and A. Lecko, The sharp bound of the third Hankel determinant for functions of bounded turning, Bol. Soc. Mat. Mex. (3) **27** (2021), no. 3, Paper No. 69, 13 pp.

[9] D. V. Krishna and T. RamReddy, Second Hankel determinant for the class of Bazilevic functions, Stud. Univ. Babeş-Bolyai Math. **60** (2015), no. 3, 413–420.

[10] S. S. Kumar and G. Kamaljeet, A cardioid domain and starlike functions, Anal. Math. Phys. **11** (2021), no. 2, Paper No. 54, 34 pp

[11] O. S. Kwon, A. Lecko and Y. J. Sim, On the fourth coefficient of functions in the Carathéodory class, Comput. Methods Funct. Theory **18** (2018), no. 2, 307–314

[12] O. S. Kwon, A. Lecko and Y. J. Sim, The bound of the Hankel determinant of the third kind for starlike functions, Bull. Malays. Math. Sci. Soc. **42** (2019), no. 2, 767–780.

[13] A. Lecko, Y. J. Sim and B. Śmiarowska, The sharp bound of the Hankel determinant of the third kind for starlike functions of order 1/2, Complex Anal. Oper. Theory **13** (2019), no. 5, 2231–2238.

[14] R. J. Libera and E. J. Złotkiewicz, Coefficient bounds for the inverse of a function with derivative in $\mathcal{P}$, Proc. Amer. Math. Soc. **87** (1983), no. 2, 251–257.

[15] R. J. Libera and E. J. Złotkiewicz, Early coefficients of the inverse of a regular convex function, Proc. Amer. Math. Soc. **85** (1982), no. 2, 225–230

[16] W. C. Ma and D. Minda, A unified treatment of some special classes of univalent functions, in *Proceedings of the Conference on Complex Analysis (Tianjin, 1992)*, 157–169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge

[17] J. W. Noonan and D. K. Thomas, On the second Hankel determinant of areally mean p-valent functions, Trans. Amer. Math. Soc. **223** (1976), 337–346.

[18] C. Pommerenke, "On the coefficients and Hankel determinants of univalent functions," J. London Math. Soc., vol. 41, pp. 111–122, 1966.

[19] C. Pommerenke, *Univalent functions*, Studia Mathematica/Mathematische Lehrbücher, Band XXV, Vandenhoeck & Ruprecht, Göttingen, 1975

[20] B. Rath et al., The Sharp Bound of the Third Hankel Determinant for Starlike Functions of Order 1/2, Complex Anal. Oper. Theory **16** (2022), no. 5, Paper No. 65, 8 pp.

[21] P. Zaprawa, Third Hankel determinants for subclasses of univalent functions, Mediterr. J. Math. **14** (2017), no. 1, Paper No. 19, 10 pp.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: nehaverma1480@gmail.com

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

Journal of Marine Science and Technology

Volume 30 | Issue 3 Article 4

A Hybrid Finite Difference Method for Singularly Perturbed Delay Partial Differential Equations with Discontinuous Coefficient and Source

Nitika Sharma
Department of Applied Mathematics, Delhi Technological University, Delhi -110042 India, nitika_phd2k18@dtu.ac.in

Aditya Kaushik
Department of Applied Mathematics, Delhi Technological University, Delhi -110042 India, akaushik@dtu.ac.in

Follow this and additional works at: https://jmstt.ntou.edu.tw/journal

Part of the Fresh Water Studies Commons, Marine Biology Commons, Ocean Engineering Commons, Oceanography Commons, and the Other Oceanography and Atmospheric Sciences and Meteorology Commons

Recommended Citation

Sharma, Nitika and Kaushik, Aditya (2022) "A Hybrid Finite Difference Method for Singularly Perturbed Delay Partial Differential Equations with Discontinuous Coefficient and Source," *Journal of Marine Science and Technology*: Vol. 30: Iss. 3, Article 4.
DOI: 10.51400/2709-6998.2578
Available at: https://jmstt.ntou.edu.tw/journal/vol30/iss3/4

This Research Article is brought to you for free and open access by Journal of Marine Science and Technology. It has been accepted for inclusion in Journal of Marine Science and Technology by an authorized editor of Journal of Marine Science and Technology.

RESEARCH ARTICLE

A Hybrid Finite Difference Method for Singularly Perturbed Delay Partial Differential Equations with Discontinuous Convection Coefficient and Source Term

Nitika Sharma [a], Aditya Kaushik [a,b,*]

[a] Department of Applied Mathematics, Delhi Technological University, Delhi 110042 India
[b] University Institute of Engineering and Technology, Panjab University, Chandigarh 160014 India

Abstract

The article presents a hybrid finite difference scheme to solve a singularly perturbed parabolic functional differential equation with discontinuous coefficient and source. The simultaneous presence of deviating argument with a discontinuous source and coefficient makes the problem stiff. The solution of the problem exhibits turning point behaviour across discontinuity as ε tends to zero. The hybrid scheme presented is a composition of a central difference scheme and a midpoint upwind scheme on a specially generated mesh. At the same time, an implicit finite difference method is used to discretize the time variable. Consistency, stability, and convergence of the presented numerical approach have been investigated. The presented method converges uniformly independent of the perturbation parameter. Numerical results have been presented for two test examples that verify the effectiveness of the scheme.

Keywords: Singular perturbation, Functional differential equation, Finite difference method, Interior layer

1. Introduction

Differential equations model a wide range of phenomena in almost every branch of sciences and engineering. Often mathematical models assume specific behaviour or phenomenon which depends on the present and the past states of a system [17,42]. In other words, previous events have a direct impact on future outcomes. Modelling such systems leads to functional differential equations that are more realistic and frequently appear in a wide range of applications. Classical examples cover oscillatory, excitable as well as chaotic behavior [1] in physiological control systems [29], nonlinear optics [19], population dynamics [16] and neuroscience [31]. Particularly in nonlinear optics, the finite-time communication delays are typically much larger than the device's internal time-scales [39], and therefore, give rise to rich dynamical phenomena [27]. Specific examples of delay systems with (multiple) large delays include semiconductor lasers with two optical feedback loops of different lengths [30], ring-cavity lasers with optical feedback [12], and others [10,37].

When we associate a mathematical model with physical phenomena, we often capture the essentials and ignore the minor components involving small parameters. A mathematical model with small parameters is called a perturbed model, while an unperturbed or reduced model is a degenerate simplified model. If a small parameter is multiplied with the highest-order derivative term, then the problem is called singularly perturbed from a mathematical perspective [4]. In that case, the

Received 6 January 2022; revised 2 March 2022; accepted 19 April 2022.
Available online 12 August 2022

* Corresponding author.
E-mail address: akaushik@dtu.ac.in (A. Kaushik).

https://doi.org/10.514/2709-6998.2578
2709-6998/© 2022 National Taiwan Ocean University.

corresponding degenerate problem cannot satisfy the prescribed boundary/initial conditions. Moreover, the stiffness, attributed to the simultaneous presence of a deviating argument and discontinuous coefficient, give rise to multiple time scales as ε approaches the limiting value zero. There exist narrow regions across the turning points where the solution varies exponentially and approaches a discontinuous limit [28,36]. In these regions, the solution reveals sharp interior layers, and it is not easy to find an asymptotic expansion in terms of ε or to find a consistent numerical approximation.

The numerical methods on a uniform mesh fail to approximate singular perturbation functional differential equations. They require an unacceptably large number of mesh points to sustain the approximation because the mesh width depends on the perturbation parameter. This limitation of the conventional numerical methods has encouraged researchers to develop robust or adaptive numerical techniques that perform well enough independent of the perturbation parameter. Many researchers have made efforts to develop parameter uniform methods for partial functional differential equations; for an overview, see [5,20] and references therein. In [2], the authors studied a Dirichlet boundary value problem for the partial functional differential equation. In this paper, a numerical method comprising a standard finite difference operator (centered in space, implicit in time) is presented. The method presented is robust in the sense that its numerical solution converges to the exact solution uniformly. The higher-order methods are of particular interest and reasonably understood for problems without a deviating argument [8,23]. Nevertheless, attempts of having a higher order of uniform convergence for parabolic functional differential equations are lacking. In [25], a higher-order uniformly convergent hybrid method is proposed for parabolic problems with a delay. A scheme of HODIE type [24] is used in spatial direction over a generalized Shishkin mesh, whereas the implicit Euler scheme discretizes the time variable. Richardson extrapolation improves the order of convergence in time-variable. It is proved that the computed solution is uniformly convergent of $\mathscr{O}((L/N)^4 + (\Delta t)^2)$. In [22], a stable finite difference scheme is presented for singularly perturbed differential-difference equations with delay and advanced term. In [13], a standard finite difference method on an equidistributed layer adapted mesh is presented. The method approximates a singularly perturbed parabolic functional differential equations of convection-diffusion type using an upwind finite difference method on a piecewise uniform Shishkin mesh. In [38], a singularly perturbed differential–difference equation with small shifts is considered. The terms containing the delay and advance parameters are approximated by using the Taylor's series expansion. A quadratic B-spline collocation method is used in the space direction on exponentially graded mesh while the Crank–Nicolson finite difference method is used in the time direction on uniform mesh.

For a time-delayed convection-diffusion problem, a hybrid scheme is reported in [9]. The method presented is a composition of an upwind scheme and a central difference scheme and has almost second-order accuracy in space which is optimal compared to [14]. While a hybrid difference scheme for parabolic differential equation with delay and advance terms is studied in [15,34]. Besides, uniformly convergent methods based on fitted mesh and fitted operator approach are presented in [6]. The fitted operator methods used have coefficients of exponential type and are quite capable of reproducing the multiscale character of the exact solution of the problem [36]. On the other hand, the fitted mesh method [32,41] adopted concentrates on the appropriate choice of the grid points in narrow layer regions. Moreover, a Robin type boundary value problem for parabolic functional differential equation is brought to attention in [3]. A numerical method comprising a standard finite difference scheme on a rectangular piecewise uniform fitted mesh is presented. It is shown that the errors are bounded in the maximum norm by $C(N_x^{-2}\ln^2 N_x + N_t^{-1})$. In [21], the authors considered a class of singularly perturbed parabolic problem with delay and discontinuous data. An implicit numerical scheme is used to find the numerical solution. The method is uniformly convergent of $\mathscr{O}(N^{-1}\ln^2 N + \Delta t)$. However, the analysis of higher-order methods for singularly perturbed parabolic partial functional differential equations with discontinuous data and degenerating convective terms has seen little development and lacks due attention. Consequently, this paper aims to extend the idea of parameter uniform difference schemes to solve a singularly perturbed parabolic functional differential equation with discontinuity in coefficients with turning points and presents a higher-order accurate approximation.

2. The continuous problem

Consider the following initial-boundary-value problem on the domain $S^- \cup S^+ = (0,1) \times (0,T] \cup (1,2) \times (0,T]$:

$$\left.\begin{aligned} &L_\varepsilon u = \varepsilon u_{xx}(x,t) + a(x)u_x(x,t) - b(x)u(x,t) - u_t(x,t) \\ &\quad = f(x,t) + c(x)u(x-1,t) \text{ in } S^- \cup S^+, \\ &u(x,t) = p_0(x) \text{ on } [0,2] \times \{t=0\}, \\ &u(x,t) = p_1(x,t) \text{ in } [-1,0] \times [0,T], \\ &u(x,t) = p_2(t) \text{ on } \{x=2\} \times [0,T], \end{aligned}\right\} \tag{2.1}$$

where $\varepsilon \ll 1$ is a small positive parameter, b and c are sufficiently smooth functions such that $c(x) > 0$, $b(x) > 0$ for all $x \in [0, 2]$. Moreover, we assume that

$$\left.\begin{aligned} &a(x) = \begin{cases} a_1(x) \text{ if } 0 \le x \le 1 \\ a_2(x) \text{ if } 1 < x \le 2, \end{cases} f(x,t) = \begin{cases} f_1(x,t) \text{ if } (x,t) \in \overline{S}^- \\ f_2(x,t) \text{ if } (x,t) \in \overline{S}^+, \end{cases} \\ &-\gamma_1^* < a_1(x) < -\gamma_1 < 0, \quad \gamma_2^* > a_2(x) > \gamma_2 > 0, \\ &|[a]| \le C, \quad |[f]| \le C, \end{aligned}\right\} \tag{2.2}$$

where $\gamma = \min\{\gamma_1, \gamma_2\}$ and $\gamma^* = \max\{\gamma_1^*, \gamma_2^*\}$. The solution of (2.1) satisfies $[u] = 0$ and $[u_x] = 0$ at $x = 1$. Here, $[u]$ denotes the jump of u defined at the point of discontinuity $x = 1$ as $[u](1, t) = u(1^+, t) - u(1^-, t)$ where $u(1^\pm, t) = \lim_{x \to 1 \pm 0} u(x,t)$.

The functions p_0, p_1 and p_2 are Hölder continuous and the compatibility conditions hold at the corners $(0,0)$, $(2,0)$ and at the transition point $(1,0)$. On the domain S^-, the delay term $u(x-1,t) = p_1(x-1,t)$. Under these assumptions, the solution of (2.1) exists and unique $u \in C^{1+\lambda}(D = (0,2) \times (0,T]) \cap C^{2+\lambda}(S^- \cup S^+)$ [26,35].

The presence of a discontinuity and a delay makes the problem challenging. The solution $u(x,t)$ of (2.1) displays a strong interior layer in the neighbourhood of the point $x = 1$. Moreover, it is easy to follow that the differential operator L_ε satisfies the following maximum principle on $\overline{D}$.

Lemma 2.1 Suppose $P \in C^0(\overline{D}) \cap C^2(S^- \cup S^+)$ satisfies $P(x,t) \le 0$ for all $(x,t) \in \Gamma := \overline{D} \setminus D$ and $L_\varepsilon P(x,t) \ge 0$ for all $(x,t) \in S^- \cup S^+$. Then $P(x,t) \le 0$ for all $(x,t) \in \overline{D}$.

Proof. Choose $(x^k, t^k) \in D$ such that $P(x^k, t^k) = \max_{(x,t) \in \overline{D}} P(x,t)$. Consequently,

$$P_x(x^k,t^k) = 0, \quad P_t(x^k,t^k) = 0 \quad \text{and} \quad P_{xx}(x^k,t^k) < 0.$$

Suppose $P(x^k, t^k) > 0$ and it follows that $(x^k, t^k) \notin \Gamma$. If $(x^k, t^k) \in S^- \cup S^+$, note that $L_\varepsilon P(x^k, t^k) < 0$. A contradiction to the assumption and consequently the required result follows.

Establishing the boundedness of solution is an important application of the maximum principle. As an immediate application, we obtain

Lemma 2.2 Let u be the solution of (2.1). Then

$$\| u \|_{\infty,\overline{D}} \le \| u \|_{\infty,\Gamma} + \frac{1}{\gamma} \| f \|_{\infty,\overline{D}}, \quad \gamma = \min\{\gamma_1, \gamma_2\}. \tag{2.3}$$

Proof. Consider

$$\psi_\pm = \begin{cases} -\| u \|_{\infty,\Gamma} - \dfrac{x}{\gamma} \| f \|_{\infty,\overline{D}} \pm u & \text{if } x \le 1 \\ -\| u \|_{\infty,\Gamma} - \dfrac{(2-x)\| f \|_{\infty,\overline{D}}}{\gamma} \pm u & \text{if } x \ge 1. \end{cases}$$

For $(x,t) \in S^-$, it follows that

$$\begin{aligned} L_\varepsilon \psi_\pm(x,t) &= \pm L_\varepsilon u - a_1(x) \frac{\| f \|}{\gamma} + b(x) \| u \| \\ &\quad + b(x) \frac{x \| f \|}{\gamma} \ge 0 \end{aligned}$$

since $a_1(x) \le 0$ and $b(x) > 0$. Similarly, for $(x,t) \in S^+$, it is easy to verify that $L_\varepsilon \psi_\pm(x,t) \ge 0$. The required result (2.3) now follows from Lemma 2.1.

In general, one can assume homogeneous boundary conditions $p_0 = p_1 = p_2 = 0$ by subtracting from u some suitable smooth function that satisfies the original boundary conditions [36]. To find sharper bounds on the solution and its derivatives, we decompose the solution u into smooth and singular components. We write $u = v + w$. The smooth component v is solution of

$$\left.\begin{aligned} &L_\varepsilon v(x,t) = f(x,t) + c(x)v(x-1,t) \text{ in } S^- \cup S^+, \\ &v(x,t) = u(x,t) \text{ in } [-1,0] \times [0,T], \\ &v(1^-,t) = j_1(t), \ v(1^+,t) = j_2(t), \ t \in (0,T], \\ &v(x,t) = u(x,t) \text{ on } \{x=2\} \times (0,T], \\ &v(x,t) = u(x,t) \text{ on } [0,2] \times \{t=0\}, \end{aligned}\right\} \tag{2.4}$$

where the functions $j_1(t)$ and $j_2(t)$ will be computed using the Theorem 2.4. The singular component w is solution of

$$\left.\begin{aligned} &L_\varepsilon w(x,t) = c(x)w(x-1,t) \text{ in } S^- \cup S^+, \\ &w(x,t) = 0 \text{ in } [-1,0] \times [0,T], \\ &w(x,t) = 0 \text{ on } \{x=2\} \times (0,T], \\ &w(x,t) = 0 \text{ on } [0,2] \times \{t=0\}, \\ &[w]\left(1,t\right) = -[v]\left(1,t\right), \\ &\left[\frac{\partial w}{\partial x}\right]\left(1,t\right) = -\left[\frac{\partial v}{\partial x}\right]\left(1,t\right), t \in \left(0,T\right] \end{aligned}\right\} \tag{2.5}$$

Following are the direct results from [18] that will be used in Theorem 2.4.

Theorem 2.3 Let $a, b, c \in C^2[0,2]$, $f \in C^{2+\lambda}(\overline{D})$ and $a(x) \geq \gamma > 0$ for $x \in [0,2]$. Also suppose p_0, p_1 and p_2 are identically zero so that the compatibility conditions hold in $\Gamma_c = (-1,0) \cup (0,0) \cup (2,0) \cup (1,0)$ and

$$\frac{\partial^{k+m} F}{\partial x^k \partial t^m} = 0, \quad 0 \leq k + 2m \leq 2,$$

where $F(x,t) = f(x,t) + c(x)u(x-1,t)$. Then $u \in C^{4+\lambda}(\overline{D})$ and

$$\left\| \frac{\partial^{k+m} u}{\partial x^k \partial t^m} \right\|_{\infty, D} \leq C\varepsilon^{-k}, \quad 0 \leq k + 2m \leq 4. \tag{2.6}$$

Theorem 2.4 There exist smooth functions $j_1(t)$, $j_2(t)$ such that the smooth and singular component defined in (2.4) and (2.5), respectively, satisfy the following bounds for $0 \leq k \leq 3$, $m \geq 0$ and $0 \leq k + 2m \leq 4$

$$\left\| \frac{\partial^{k+m} v}{\partial x^k \partial t^m} \right\|_{\infty, S^- \cup S^+} \leq C, \quad \left\| \frac{\partial^4 v}{\partial x^4} \right\|_{\infty, S^- \cup S^+} \leq C\varepsilon^{-1}$$

and

$$\left| \frac{\partial^{k+m} w}{\partial x^k \partial t^m}(x,t) \right| \leq \begin{cases} C\left(\varepsilon^{-k} \exp(-(1-x)\gamma_1/\varepsilon)\right) & \text{for } (x,t) \in S^- \\ C\left(\varepsilon^{-k} \exp(-(x-1)\gamma_2/\varepsilon)\right) & \text{for } (x,t) \in S^+, \end{cases}$$

$$\left| \frac{\partial^4 w}{\partial x^4}(x,t) \right| \leq \begin{cases} C\left(\varepsilon^{-4} \exp(-(1-x)\gamma_1/\varepsilon)\right) & \text{for } (x,t) \in S^- \\ C\left(\varepsilon^{-4} \exp(-(x-1)\gamma_2/\varepsilon)\right) & \text{for } (x,t) \in S^+. \end{cases}$$

Proof. We conduct our analysis separately in the subregions S^- and S^+ to obtain stronger bounds on the solution and its derivatives. We start with the subregion S^+ and define $\overline{D}^*$ such that $\overline{S}^+ \subset \overline{D}^*$. Let $D^* = \Omega^* \times (0,T]$, $\Omega^* = [-2,2]$ and define a function v^* on D^* as

$$v^* = v_0^* + \varepsilon v_1^* + \varepsilon^2 v_2^* + \varepsilon^3 v_3^*,$$

where the functions v_0^*, v_1^* and v_2^* are solutions of

$$\left.\begin{aligned} & a^* \frac{\partial v_0^*}{\partial x} - b^* v_0^* - \frac{\partial v_0^*}{\partial t} = f^* + c^* v_0^*(x-1,t) \quad \text{in} \quad D^*, \\ & v_0^*(x,t) = p_1^*(x,t) \quad \text{in} \quad [-3,0] \times [0,T], \\ & v_0^*(x,t) = p_0^*(x) \quad \text{on} \quad [-2,2] \times \{t=0\}, \\ & a^* \frac{\partial v_i^*}{\partial x} - b^* v_i^* - \frac{\partial v_i^*}{\partial t} = -\frac{\partial^2 v_{i-1}^*}{\partial x^2} + c^* v_i^*(x-1,t) \quad \text{in} \quad D^*, \\ & v_i^*(x,t) = 0 \quad \text{in} \quad [-3,0] \times (0,T], v_i^*(x,0) = 0 \quad \text{on} \quad [-2,2] \times \{t=0\}, \quad i = 1,2 \end{aligned}\right\} \tag{2.7}$$

and v_3^* is solution of

$$\left.\begin{aligned} & \varepsilon \frac{\partial^2 v_3^*}{\partial x^2} + a^* \frac{\partial v_3^*}{\partial x} - b^* v_3^* - \frac{\partial v_3^*}{\partial t} = c^* v_3^*(x-1,t) - \frac{\partial^2 v_2^*}{\partial x^2} \text{ in } D^*, \\ & v_3^*(x,t) = 0 \text{ in } [-3,0] \times [0,T], \\ & v_3^*(x,0) = 0 \text{ on } [-2,2] \times \{t=0\}, \\ & v_3^*(x,t) = 0 \text{ on } \{x=2\} \times (0,T] \end{aligned}\right\} \tag{2.8}$$

Here, the coefficients a^*, b^* and c^* as well as the condition p_0^* are respective smooth extensions of a, b, c and p_0 from the domain $[1,2]$ to the domain $[-2,2]$. The functions f^* and p_1^* are the smooth extensions of f and p_1 from the domain $\overline{S}^+$ to the domain $\overline{D}^*$. In a neighbourhood of the point $(-2,0)$, the functions p_0^*, p_1^* and f^* are built such that $p_0^* = p_1^* = f^* = 0$. Assuming that a^*, b^*, c^* and f^* are sufficiently smooth on $\overline{D}^*$. We set all the initial-boundary data associated with (2.7) equal to zero. Define $F^* = f^* + c^* v^*(x-1,t)$ and impose the following compatibility condition on the set Γ_c:

$$\frac{\partial^{k+m} F^*}{\partial x^k \partial t^m} = 0 \quad \text{for} \quad 0 \leq k + m \leq 7. \tag{2.9}$$

Follows from the results of [7] for the first order differential equations defined in (2.7), we have $v_i^* \in C^{9-2i+\lambda}(\overline{D}^*) \cap C^{8-2i}(\overline{D}^*)$, $i = 0,1,2$ which implies $\frac{\partial^2 v_2^*}{\partial x^2} \in C^{2+\lambda}(\overline{D}^*)$ and therefore $v_3^* \in C^{4+\lambda}(\overline{D}^*)$. Then it follows that

$$\left\| \frac{\partial^{k+m} v_i^*}{\partial x^k \partial t^m} \right\|_{\infty, D^*} \leq C, \quad i = 0,1,2,$$

$$\left\| \frac{\partial^{k+m} v_3^*}{\partial x^k \partial t^m} \right\|_{\infty, D^*} \leq C\varepsilon^{-k} \quad \text{for} \quad 0 \leq k + 2m \leq 4.$$

The smooth component v is now defined as a restriction of v^* on the domain $\overline{S}^+$. Define $v^*(1,t) = j_2(t) = v(1,t)$. Thus v satisfies

$$L_\varepsilon v = f + c(x)v(x-1,t) \quad \text{in} \quad S^+,$$

$$v(x,t) = v^*(x,t) \quad \text{on} \quad \{x=1\}\times(0,T],$$

$$v(x,t) = u(x,t) \quad \text{on} \quad [1,2]\times\{t=0\}.$$

Since $v^* = v_0^* + \varepsilon v_1^* + \varepsilon^2 v_2^* + \varepsilon^3 v_3^*$, thus

$$\left\|\frac{\partial^{k+m}v^*}{\partial x^k\partial t^m}\right\|_{\infty,S^+} \le \left\|\frac{\partial^{k+m}v_0^*}{\partial x^k\partial t^m}\right\|_{\infty,S^+} + \varepsilon\left\|\frac{\partial^{k+m}v_1^*}{\partial x^k\partial t^m}\right\|_{\infty,S^+} + \varepsilon^2\left\|\frac{\partial^{k+m}v_2^*}{\partial x^k\partial t^m}\right\|_{\infty,S^+} + \varepsilon^3\left\|\frac{\partial^{k+m}v_3^*}{\partial x^k\partial t^m}\right\|_{\infty,S^+}$$
$$\le C\left(1+\varepsilon^{3-k}\right) \quad \text{for} \quad 0\le k+2m\le 4.$$

As v is a restriction of v^* on the domain $\overline{S}^+$, we have

$$\left\|\frac{\partial^{k+m}v}{\partial x^k\partial t^m}\right\|_{\infty,S^+} \le C\left(1+\varepsilon^{3-k}\right) \quad \text{for} \quad 0\le k+2m\le 4.$$

Thus for $0 \le k \le 3$ and $0 \le k+2m \le 4$, the smooth component v satisfies

$$\left\|\frac{\partial^{k+m}v}{\partial x^k\partial t^m}\right\|_{\infty,S^+} \le C \quad \text{and} \quad \left\|\frac{\partial^4 v}{\partial x^4}\right\|_{\infty,S^+} \le C\varepsilon^{-1}.$$

Similarly, the desired bounds for smooth component v in the subregion S^- can be obtained. Define the barrier functions on $S^-\cup S^+$

$$\Phi^\pm(x,t) = \begin{cases} (\pm v - C)\exp\left(\dfrac{-(1-x)\gamma_1}{\epsilon}\right) \pm w, & (x,t)\in S^- \\ (\pm v - C)\exp\left(\dfrac{-(x-1)\gamma_2}{\epsilon}\right) \pm w, & (x,t)\in S^+, \end{cases} \tag{2.10}$$

where C is a constant. For $(x,t)\in S^-$, using assumption $(a+\gamma_1)<0$ to obtain

$$L_\varepsilon\phi^\pm(x,t) = L_\varepsilon\left((\pm v - C)\exp\left(\frac{-(1-x)\gamma_1}{\varepsilon}\right)\right) \pm L_\varepsilon w(x,t) \ge 0. \tag{2.11}$$

Similarly, for $(x,t)\in S^+$ we obtain

$$L_\varepsilon\phi^\pm(x,t) = \frac{\gamma_2}{\varepsilon}\exp\left(\frac{-(x-1)\gamma_2}{\varepsilon}\right)\left((\pm v - C)(-a+\gamma_2) - 2\varepsilon v_x + \frac{Cb\varepsilon}{\gamma_2}\right)$$
$$\pm\frac{\gamma_2}{\varepsilon}\exp\left(\frac{-(x-1)\gamma_2}{\varepsilon}\right)\left(\frac{f + cv(x-1,t)}{\gamma_2}\right) \ge 0. \tag{2.12}$$

Moreover, $\phi^\pm\in C^0(\overline{D})$ and $\phi^\pm(x,t)\le 0$ for $(x,t)\in\Gamma$. Consequently, the required bounds on w follows from Lemma 2.1. The required estimates for the derivatives of w follow from [33].

3. Difference scheme

The solution of the problem exhibits strong interior layer at $x = 1$. Therefore, we descretize the domain by constructing a rectangular mesh $\overline{D}^{N,M} = \overline{S}_x^N \times T_t^M$ in such a way that it will condense around the point $x = 1$. We write

$$[0,2] = [0,1-\sigma]\cup[1-\sigma,1]\cup[1,1+\sigma]\cup[1+\sigma,2],$$

where $\sigma = \min\{\frac{1}{2}, \sigma_0\varepsilon\ln N\}$ and σ_0 is a constant that will be chosen later on. We place $\frac{N}{4}$ mesh points in each of the subintervals. Consequently, we obtain

$$S_x^N = \left\{x_i : i = 1,2\ldots,\frac{N}{2}-1\right\}\cup\left\{x_i : i = \frac{N}{2}+1,\ldots,N-1\right\}$$

and

$$h_i = x_i - x_{i-1}, \quad i = 1,2,\ldots,N,$$
$$\widehat{h}_i = h_i + h_{i+1}, \; i = 1,2,\ldots,N-1,$$

We define the uniform mesh for the domain $[0,T]$, as follows

$$T_t^M = \{t_k = k\Delta t, k = 0, \ \ldots \ , M, \Delta t = T/M\}.$$

$$h_i = \begin{cases} H = \dfrac{4(1-\sigma)}{N}, & i = 1,\ldots,N/4, 3N/4+1,\ldots,N \\ h = \dfrac{4\sigma}{N}, & i = N/4+1,\ldots,N/2, N/2+1,\ldots,3N/4. \end{cases}$$

To discretize the differential operator in (2.1), we first define the finite difference operators on the mesh $\overline{D}^{N,M}$ as

$$D_x^+ v_i^k = \frac{v_{i+1}^k - v_i^k}{h_{i+1}}, \quad D_x^- v_i^k = \frac{v_i^k - v_{i-1}^k}{h_i},$$
$$D_x^0 v_i^k = \frac{v_{i+1}^k - v_{i-1}^k}{h_{i+1}+h_i}, \; D_t^- v_i^k = \frac{v_i^k - v_i^{k-1}}{\Delta t},$$

and $\delta_x^2 v_i^k = \frac{2(D_x^+ v_i^k - D_x^- v_i^k)}{h_{i+1}+h_i}$. Also define $v_{i\mp\frac{1}{2}}^k = \frac{(v_{i\mp1}^k + v_i^k)}{2}$, $v_{i\mp\frac{1}{2}} = \frac{(v_{i\mp1}+v_i)}{2}$.

We use the classical central difference scheme in the intervals $(1-\sigma,1),(1,1+\sigma)$ and the midpoint upwind scheme in the rest of the intervals. At the point of discontinuity, second-order one-sided difference approximations are used to keep the continuity of the spatial derivative. We use the backward-Euler method for discretizing the time derivative. The discrete problem thus reads:

$$\begin{cases} U_{i,0} = p_0(x_i) & \text{for } i=0,\ldots,N \\ L_{mu}^{N,M,(L)} U_{i,k+1} = f_{\frac{i-1}{2},k+1} + c_{\frac{i-1}{2}} p_1 & \text{for } i=1,\ldots,\dfrac{N}{4} \\ L_{cen}^{N,M} U_{i,k+1} = f_{i,k+1} + c_i p_1 & \text{for } i=\dfrac{N}{4}+1,\ldots,\dfrac{N}{2}-1 \\ L_{cen}^{N,M} U_{i,k+1} = f_{i,k+1} + c_i U_{i-N/2,k+1} & \text{for } i=\dfrac{N}{2}+1,\ldots,\dfrac{3N}{4}-1 \\ L_{mu}^{N,M,(R)} U_{i,k+1} = f_{\frac{i+1}{2},k+1} + c_{\frac{i+1}{2}} U_{i-N/2,k+1} & \text{for } i=\dfrac{3N}{4},\ldots,N-1 \\ D_x^F U_{i,k+1} - D_x^B U_{i,k+1} = 0 & \text{for } i=\dfrac{N}{2} \\ \text{for } k=0,\ldots,M-1, \end{cases} \tag{3.1}$$

where

$$\begin{cases} \mathcal{L}_{mu}^{N,M,(L)} U_{i,k+1} &= \epsilon\delta_x^2 U_{i,k+1} + a_{\frac{i-1}{2}} D_x^- U_{i,k+1} - b_{\frac{i-1}{2}} U_{\frac{i-1}{2},k+1} - D_t^- U_{\frac{i-1}{2},k+1} \\ L_{cen}^{N,M} U_{i,k+1} &= \epsilon\delta_x^2 U_{i,k+1} + a_i D_x^0 U_{i,k+1} - b_i U_{i,k+1} - D_t^- U_{i,k+1} \\ \mathcal{L}_{mu}^{N,M,(R)} U_{i,k+1} &= \epsilon\delta_x^2 U_{i,k+1} + a_{\frac{i+1}{2}} D_x^+ U_{i,k+1} - b_{\frac{i+1}{2}} U_{\frac{i+1}{2},k+1} - D_t^- U_{\frac{i+1}{2},k+1} \end{cases} \tag{3.2}$$

and

$$\begin{cases} D_x^F U_{\frac{N}{2},k} = \left(-U_{\frac{N}{2}+2,k+1} + 4U_{\frac{N}{2}+1,k+1} - 3U_{\frac{N}{2},k+1} \right) \Big/ 2h \\ D_x^B U_{\frac{N}{2},k} = \left(U_{\frac{N}{2}-2,k+1} - 4U_{\frac{N}{2}-1,k+1} + 3U_{\frac{N}{2},k+1} \right) \Big/ 2h. \end{cases} \tag{3.3}$$

After simplifying the terms in (3.1), we obtain

$$\begin{cases} U_i^0 = p_0(x_i) & \text{for} \quad i = 0, \ldots, N \\ L_\varepsilon^{N,M} U_{i,k+1} = \tilde{f}_{i,k+1} & \text{for} \quad i = 1, \ldots, N-1 \\ U_{i,k+1} = p_1(x_i, t_{k+1}) & \text{for} \quad i = -N/2, \ldots, 0 \\ U_{N,k+1} = p_2(t_{k+1}) & \text{for} \quad k = 0, \ldots, M-1, \end{cases} \tag{3.4}$$

where

$$L_\epsilon^{N,M} U_{i,k+1} = \begin{cases} \left[r_i^- U_{i-1,k+1} + r_i^0 U_{i,k+1} + r_i^+ U_{i+1,k+1} \right] + \left[p_i^- U_{i-1,k} + p_i^0 U_{i,k} + p_i^+ U_{i+1,+k+1} \right] \\ \text{for} \quad i = 1, \ldots, N/2-1, N/2+1, \ldots, N-1 \\ q_i^{-,2} U_{i-2,k+1} + q_i^{-,1} U_{i-1,k+1} + q_i^0 U_{i,k+1} + q_i^{+,1} U_{i+1,k+1} + q_i^{+,2} U_{i+2,k+1} \\ \text{for} \quad i = N/2 \end{cases} \tag{3.5}$$

and

$$\tilde{f}_{i,k+1} = \begin{cases} m_i^- f_{i-1,k+1} + m_i^0 f_{i,k+1} + m_i^+ f_{i+1,k+1} + s_i^+ p_1(x_i, t_{k+1}) + s_i^- U_{i-N/2,k+1} \\ \text{for} \quad i = 1, \ldots, N/2-1, N/2+1, \ldots, N-1 \\ 0 \quad \text{for} \quad i = N/2. \end{cases} \tag{3.6}$$

Here, for $i \in \{1 : N/4\}$

$$\begin{cases} r_i^- = \left(\dfrac{2\epsilon}{\widehat{h}_i h_i} - \dfrac{a_{i-1/2}}{h_i} - \dfrac{b_{i-1/2}}{2} - \dfrac{1}{2\Delta t}\right), \\ r_i^0 = \left(\dfrac{-2\epsilon}{h_i h_{i+1}} + \dfrac{a_{i-1/2}}{h_i} - \dfrac{b_{i-1/2}}{2} - \dfrac{1}{2\Delta t}\right), \\ r_i^+ = \dfrac{2\epsilon}{\widehat{h}_i h_{i+1}}, \quad p_i^- = \dfrac{1}{2\Delta t}, \quad p_i^0 = \dfrac{1}{2\Delta t}, \\ p_i^+ = 0, \quad m_i^- = \dfrac{1}{2}, \quad m_i^0 = \dfrac{1}{2}, \quad m_i^+ = 0, \\ s_i^+ = c_{i-1/2}, \quad s_i^- = 0, \end{cases} \tag{3.7}$$

for $i \in \{N/4 + 1 : 3N/4 - 1\}/\{\frac{N}{2}\}$

$$\begin{cases} r_i^- = \left(\dfrac{2\varepsilon}{\widehat{h}_i h_i} - \dfrac{a_i}{\widehat{h}_i}\right), \quad r_i^0 = \left(\dfrac{-2\epsilon}{h_i h_{i+1}} - b_i - \dfrac{1}{\Delta t}\right), \\ r_i^+ = \left(\dfrac{2\epsilon}{\widehat{h}_i h_{i+1}} + \dfrac{a_i}{\widehat{h}_i}\right), \quad p_i^- = 0, \quad p_i^0 = \dfrac{1}{\Delta t}, \\ p_i^+ = 0, \quad m_i^- = 0, \quad m_i^0 = 1, \quad m_i^+ = 0, \end{cases} \tag{3.8}$$

for $i \in \{N/4 + 1 : N/2 - 1\}$

$$s_i^+ = c_i, \quad s_i^- = 0,$$

for $i \in \{N/2 + 1 : 3N/4 - 1\}$

$$s_i^+ = 0, \quad s_i^- = c_i,$$

for $i \in \{3N/4 : N - 1\}$

$$\begin{cases} r_i^- = \dfrac{2\varepsilon}{\widehat{h}_i h_i}, \quad r_i^0 = \left(\dfrac{-2\varepsilon}{h_i h_{i+1}} - \dfrac{a_{i+1/2}}{h_{i+1}} - \dfrac{b_{i+1/2}}{2} - \dfrac{1}{2\Delta t}\right), \\ r_i^+ = \left(\dfrac{2\varepsilon}{\widehat{h}_i h_{i+1}} + \dfrac{a_{i+1/2}}{\widehat{h}_{i+1}} - \dfrac{b_{i+1/2}}{2} - \dfrac{1}{2\Delta t}\right), \\ p_i^- = 0, \; p_i^0 = \dfrac{1}{2\Delta t}, \; p_i^+ = \dfrac{1}{2\Delta t}, \\ m_i^- = 0, \quad m_i^0 = \dfrac{1}{2}, \quad m_i^+ = \dfrac{1}{2}, \quad s^+ = 0, \quad s^- = c_{i+1/2} \end{cases} \tag{3.9}$$

and lastly,

$$q_{N/2}^{-,2} = \frac{-1}{2h}, \quad q_{N/2}^{-,1} = \frac{2}{h}, \quad q_{N/2}^0 = -\frac{3}{h}, \quad q_{N/2}^{+,1} = \frac{2}{h}, \quad q_{N/2}^{+,2} = \frac{-1}{2h}. \tag{3.10}$$

4. Error estimates

The difference operator $L_\varepsilon^{N,M}$ in (3.5) fails to fulfill the conditions of discrete maximum principle (as for this we need $q_i \geq 0$ to prove A to be an M-matrix defined in Equation (4.9)). As a consequence, we have to modify the equation (3.4), for $i = N/2$

$$\begin{aligned} &q_{N/2}^{-,2} U_{N/2-2,k+1} + q_{N/2}^{-,1} U_{N/2-1,k+1} + q_{N/2}^0 U_{N/2,k+1} \\ &\quad + q_{N/2}^{+,1} U_{N/2+1,k+1} + q_{N/2}^{+,2} U_{N/2+2,k+1} = 0. \end{aligned} \tag{4.1}$$

From (3.4), for $i = N/2 - 1$, we have

$$\begin{aligned} &\left(\frac{2\varepsilon - h a_{N/2-1}}{2h^2}\right) U_{N/2-2,k+1} = f_{N/2-1,k+1} \\ &+ c_{N/2-1} p_1(x_{-1}, t_{k+1}) \\ &- r_{N/2-1}^0 U_{N/2-1,k+1} + r_{N/2-1}^+ U_{N/2,k+1} \\ &+ \frac{1}{\Delta t} U_{N/2-1,k} \end{aligned} \tag{4.2}$$

and, for $i = N/2 + 1$

$$\begin{aligned} &\left(\frac{2\varepsilon + h a_{N/2+1}}{2h^2}\right) U_{N/2+2,k+1} = f_{N/2+1,k+1} + c_{N/2+1} U(x_1, t_{k+1}) \\ &- r_{N/2+1}^0 U_{N/2+1,k+1} + r_{N/2+1}^+ U_{N/2,k+1} \\ &+ \frac{1}{\Delta t} U_{N/2+1,k}. \end{aligned} \tag{4.3}$$

Now putting the values of $U_{N/2-2,k+1}$ and $U_{N/2+2,k+1}$ from Equations (4.2) and (4.3) into (4.1), we get

$$\begin{aligned} &q_{N/2}^{-,2} \left(\frac{2h^2}{2\varepsilon - h a_{N/2-1}}\right) \Big(f_{N/2-1,k+1} + c_{N/2-1} p_1 \\ &\left(x_{-1}, t_{k+1}\right) - r_{N/2-1}^0 U_{N/2-1,k+1} \\ &- r_{N/2-1}^+ U_{N/2,k+1} - \frac{1}{\Delta t} U_{N/2-1,k} \Big) + q_{N/2}^{-,1} U_{N/2-1,k+1} \\ &+ q_{N/2}^0 U_{N/2,k+1} + q_{N/2}^{+,1} U_{N/2+1,k+1} \\ &+ q_{N/2}^{+,2} \left(\frac{2h^2}{2\varepsilon + h a_{N/2+1}}\right) \Big(f_{N/2+1,k+1} + c_{N/2+1} U\Big(x_1, t_{k+1}\Big) \\ &- r_{N/2+1}^0 U_{N/2+1,k+1} \end{aligned}$$

$$-r^{+}_{N/2+1}U_{N/2,k+1}-\frac{1}{\Delta t}U_{N/2+1,k}\Big)=0. \tag{4.4}$$

After rearranging the terms in (4.4), the discrete problem thus reads

$$\begin{cases} U_{i,0}=p_0(x_i) & \text{for} \quad i=0,\ldots,N \\ L^{N,M}_{\tau}U_{i,k+1}=\tilde{f}_{\tau,i,k+1} & \text{for} \quad i=1,\ldots,N-1 \\ U_{i,k+1}=p_1(x_i,t_{k+1}) & \text{for} \quad i=-N/2,\ldots,0 \\ U_{N,k+1}=p_2(t_{k+1}) & \text{for} \quad k=0,\ldots,M-1, \end{cases} \tag{4.5}$$

where

$$L^{N,M}_{\tau}U_{i,k+1}=\begin{cases} \left(r^{-}_i U_{i-1,k+1}+r^{0}_i U_{i,k+1}+r^{+}_i U_{i+1,k+1}\right) & \\ +\left(p^{-}_i U^{k}_{i-1}+p^{0}_i U^{k}_{i}+p^{+}_i U_{i+1,k}\right) & if\, i=N/2 \\ L^{N,M}_{\epsilon}U_{i,k+1} & if\, i\neq N/2 \end{cases} \tag{4.6}$$

and

$$\tilde{f}_{\tau,i,k+1}=\begin{cases} \left[m^{-}_i f_{i-1,k+1}+m^{0}_i f_{i,k+1}+m^{+}_i f_{i+1,k+1}\right] & \\ +l_1 p_1(x_{-1},t_{k+1})+l_2 U(x_1,t_{k+1}) & \text{if } i=N/2 \\ \tilde{f}_{i,k+1} & \text{if } i\neq N/2, \end{cases}$$

where, for $i = N/2$

$$\begin{cases} r^{-}_i=\dfrac{1}{2h}\left(4-\dfrac{2\left(2\epsilon+h^2 b_{i-1}+\dfrac{h^2}{\Delta t}\right)}{2\epsilon-ha_{i-1}}\right), \\ r^{0}_i=\dfrac{1}{2h}\left(-6+\dfrac{2\epsilon+ha_{i-1}}{2\epsilon-ha_{i-1}}+\dfrac{2\epsilon-ha_{i+1}}{2\epsilon+ha_{i+1}}\right), \\ r^{+}_i=\dfrac{1}{2h}\left(4-\dfrac{2\left(2\epsilon+h^2 b_{i+1}+\dfrac{h^2}{\Delta t}\right)}{2\epsilon+ha_{i+1}}\right), \\ p^{-}_i=\dfrac{h}{(2\epsilon-ha_{i-1})\Delta t}, \quad p^{0}_i=0, \quad p^{+}_i=\dfrac{h}{(2\epsilon+ha_{i+1})\Delta t}, \\ m^{-}_i=\dfrac{h}{(2\epsilon-ha_{i-1})}, \quad m^{0}_i=0, \quad m^{+}_i=\dfrac{h}{(2\epsilon+ha_{i+1})}, \\ l_1=\dfrac{-hc_{i-1}}{2\epsilon-ha_{i-1}}, \quad l_2=\dfrac{-hc_{i+1}}{2\epsilon+ha_{i+1}} \end{cases}$$

and for $i\neq N/2$, these coefficients are defined in (3.7−3.9). Let $D^{N,M}=\overline{D}^{N,M}\cap D$ and $\Gamma=\overline{D}^{N,M}\setminus D^{N,M}$.

Lemma 4.1 Let $N\geq N_0$ be such that

$$\frac{N_0}{\ln N_0}\geq 2\sigma_0\gamma^{*} \quad \text{and} \tag{4.7}$$

$$\left(\|b\|_{\infty}+\Delta t^{-1}\right)\leq\frac{\gamma N_0}{2}. \tag{4.8}$$

Let Y be the mesh function such that $Y\leq 0$ on $\Gamma^{N,M}$ and $L^{N,M}_{\tau}Y\geq 0$ in $D^{N,M}$. Then $Y\leq 0$ in $\overline{D}^{N,M}$.

Proof. Write $L^{N,M}_{\tau}$ as

$$-L^{N,M}_{\tau}Y_{i,k+1}=\left[A_{i,i-1}Y_{i-1,k+1}+A_{i,i}Y_{i,k+1}+A_{i,i+1}Y_{i+1,k+1}\right]$$
$$-\left[B_{i,i-1}Y_{i-1,k}+B_{i,i}Y_{i,k}+B_{i,i+1}Y_{i+1,k}\right], \tag{4.9}$$

where $A:=(A_{i,j})$ and $B:=(B_{i,j})$ are written as

$$A_{i,i-1}=-r^{-}_i, \quad A_{i,i}=-r^{0}_i, \quad A_{i,i+1}=-r^{+}_i,$$

$$B_{i,i-1}=p_i^-,\quad B_{i,i}=p_i^0,\quad B_{i,i+1}=p_i^+.$$

Clearly $B\geq 0$, since $p_i^-,p_i^0,p_i^+\geq 0$ and using (4.7) and (4.8) to show A is an M-matrix [11]. Remaining part is proved by induction. Assume

$$Y_{i,k}\leq 0,\quad k=0,\ldots,N-1.$$

Then, we can rewrite (4.9) as

$$AY_{i,k+1}=BY_{i,k}-L^{N,M}Y_{i,k+1}.$$

Since $B\geq 0$, $Y_{i,k}\leq 0$, $A^{-1}\geq 0$ and $L_\tau^{N,M}Y_{i,k+1}\geq 0$ by hypothesis. Then, $Y_{i,k+1}\leq 0$ in $D^{N,M}$.

We obtain the following estimate as an immediate consequence of Lemma 4.1.

Lemma 4.2 Let U be the solution of (4.5) and the conditions (4.7) and (4.8) hold true. Then

$$\|U\|_{\infty,\overline{D}^{N,M}}\leq \|U\|_{\infty,I^{N,M}}+\frac{1}{\gamma}\|\tilde{f}_\tau\|_{\infty,\overline{D}^{N,M}}$$

Proof. Let

$$\phi_i^{\pm,k}=-\|U\|_{\infty,I^{N,M}}-\begin{cases}x_i\dfrac{\|\tilde{f}_\tau\|_\infty}{\gamma}\mp U_{i,k} & for\ 1\leq i\leq N/2\\ (2-x_i)\dfrac{\|\tilde{f}_\tau\|_\infty}{\gamma}\mp U_{i,k} & for\ N/2<i\leq N\end{cases}.$$

Then $\phi_N^{\pm,k+1}\leq 0$, $\phi_i^{\pm,k+1}\leq 0$, $-N/2\leq i\leq 0$ and $\phi_i^{\pm,0}\leq 0$, $0\leq i\leq N$. For $i\neq N/2$, $L_\tau^{N,M}\phi_i^{\pm,k+1}\geq 0$. Further

$$L_\tau^{N,M}\phi_{N/2}^{\pm,k+1}=L_\tau^{N,M}\left\{-\|U\|_{\infty,I^{N,M}}-\frac{1}{\gamma}\|\tilde{f}_\tau\|_{\infty,\overline{D}^{N,M}}\mp U_{N/2,k+1}\right\}\geq (D_x^F-D_x^B)\phi_{N/2}^{\pm,k+1}$$

and

$$(D_x^F-D_x^B)\phi_{N/2}^{\pm,k+1}=\frac{1}{2h}\left(-\phi_{N/2+2}+4\phi_{N/2+1}-6\phi_{N/2}-\phi_{N/2-2}+4\phi_{N/2-1}\right)\geq 0.$$

Consequently, the required result follows from Lemma 4.1.

Next, we decompose the solution into smooth and singular component. We write $U_{i,k+1}=V_{i,k+1}+W_{i,k+1}$. Here the mesh functions V_L and V_R satisfy

$$\left.\begin{aligned}&L_\tau^{N,M}V_{L,i,k+1}=\tilde{f}_{\tau,i,k+1}\ \text{ for }\ i=1,\ldots,N/2-1,\\&V_{L,i,k+1}=v(x_i,t_{k+1})\ \text{ for }\ i=-N/2,\ldots,0,\\&V_{L,N/2,k+1}=v(1^-,t_{k+1}),\quad k\geq 0,\\&V_{L,i,0}=v(x_i,0)\ \text{ for }\ i=1,\ldots,N/2\end{aligned}\right\}$$

and

$$\left.\begin{aligned}&L_\tau^{N,M}V_{R,i,k+1}=\tilde{f}_{\tau,i,k+1}\quad \text{for}\quad i=N/2+1,\ldots,N-1,\\&V_{R,N/2,k+1}=v(1^+,t_{k+1}),\quad V_{R,N,k+1}=v(2,t_{k+1}),\quad k\geq 0,\\&V_{R,i,0}=v(x_i,0)\quad \text{for}\quad i=N/2,\ldots,N.\end{aligned}\right\}\tag{4.11}$$

The mesh function W_L and W_R satisfy

$$\left.\begin{aligned}&L_\tau^{N,M}W_{L,i,k+1}=0\quad \text{for}\quad i=1,\ldots,N/2-1,\\&L_\tau^{N,M}W_{R,i,k+1}=0\quad \text{for}\quad i=N/2+1,\ldots,N-1,\\&W_{L,i,k+1}=0\quad \text{for}\quad i=-N/2,\ldots,0,\\&W_{R,N,k+1}=0,\quad k\geq 0,\\&W_{L,i,0}=0\quad \text{for}\quad i=0,\ldots,N/2,\\&W_{R,i,0}=0\quad \text{for}\quad i=N/2,\ldots,N,\\&W_{R,N/2,k+1}+V_{R,N/2,k+1}=W_{L,N/2,k+1}+V_{L,N/2,k+1},\\&D_x^FW_{R,N/2,k+1}+D_x^FV_{R,N/2,k+1}=D_x^BW_{L,N/2,k+1}\\&+D_x^BV_{L,N/2,k+1},k\geq 0.\end{aligned}\right\}\tag{4.12}$$

Numerical solution U satisfies

$$U_{i,k+1}=\begin{cases}V_{L,i,k+1}+W_{L,i,k+1}, & 0\leq i\leq N/2-1\\ V_{L,i,k+1}+W_{L,i,k+1}=V_{R,i,k+1}+W_{R,i,k+1}, & i=N/2\\ V_{R,i,k+1}+W_{R,i,k+1}, & N/2+1\leq i\leq N-1.\end{cases}\tag{4.13}$$

Lemma 4.3 Let V_L and V_R are the solutions of (4.10), (4.11) and v is the solution of (2.4). Then under the assumptions (4.7) and (4.8), we have

$$\begin{cases} |V_{L,i,k+1} - v(x_i, t_{k+1})| \le C\left(N^{-2} + \Delta t\right)x_i & \text{for } i = 1, \ldots, N/2 - 1 \\ |V_{R,i,k+1} - v(x_i, t_{k+1})| \le C\left(N^{-2} + \Delta t\right)(2 - x_i) & \text{for } i = N/2 + 1, \ldots, N - 1. \end{cases}$$

Proof. Consider $\Psi_{L,i}^k = -C(N^{-2} + \Delta t)x_i, \quad i \in \{0, \ldots, N/2\}$. Then

$$L_\tau^{N,M}(V_{L,i,k+1} - v(x_i, t_{k+1})) = \left(L_\varepsilon - L_\tau^{N,M}\right)v(x_i, t_{k+1}).$$

We observe that for $i \in \{1, \ldots, N/2 - 1\}$

$$\left|L_\tau^{N,M}(V_{L,i,k+1} - v(x_i, t_{k+1}))\right|$$

$$\le \begin{cases} C\left[(\epsilon + h_i)(h_i + h_{i+1})\left\|\dfrac{\partial^3 v}{\partial x^3}\right\|_\infty + h_i^2\left(\left\|\dfrac{\partial^2 v}{\partial x^2}\right\|_\infty + \left\|\dfrac{\partial v}{\partial x}\right\|_\infty\right) + \Delta t\left\|\dfrac{\partial^2 v}{\partial t^2}\right\|_\infty\right], \ 1 \le i \le N/4 \\ C\left[h^2\left(\epsilon\left\|\dfrac{\partial^4 v}{\partial x^4}\right\|_\infty + \left\|\dfrac{\partial^3 v}{\partial x^3}\right\|_\infty\right) + \Delta t\left\|\dfrac{\partial^2 v}{\partial t^2}\right\|_\infty\right], \quad N/4 + 1 \le i \le N/2 - 1. \end{cases}$$

Using Theorem 2.4, conditions $h_i \le CN^{-1}$ and $\varepsilon \le N^{-1}$ to obtain

$$\left|L_\tau^{N,M}(V_{L,i,k+1} - v(x_i, t_{k+1}))\right| \le C\left(N^{-2} + \Delta t\right) \le L_\tau^{N,M}\psi_{L,i}^{k+1}.$$

Further the required result follows from Lemma 4.1

$$|V_{L,i,k+1} - v(x_i, t_n)| \le C\left(N^{-2} + \Delta t\right)x_i, \quad 1 \le i \le N/2 - 1.$$

Consider

$$\psi_{R,i} = -C\left(N^{-2} + \Delta t\right)(2 - x_i), \quad i \in \{N/2, \ldots, N\}.$$

Similarly, for $i \in \{N/2 + 1, \ldots, N - 1\}$ we obtain

$$\left|L_\varepsilon^{N,M}(V_{R,i,k+1} - v(x_i, t_{k+1}))\right|$$

$$\le \begin{cases} C\left[h^2\left(\varepsilon\left\|\dfrac{\partial^4 v}{\partial x^4}\right\|_\infty + \left\|\dfrac{\partial^3 v}{\partial x^3}\right\|_\infty\right) + \Delta t\left\|\dfrac{\partial^2 v}{\partial t^2}\right\|_\infty\right], \quad N/2 + 1 \le i \le 3N/4 - 1 \\ C\left[(\varepsilon + h_{i+1})(h_i + h_{i+1}) \left\| \dfrac{\partial^3 v}{\partial x^3}\right\|_\infty + h_{i+1}^2\left(\left\|\dfrac{\partial^2 v}{\partial x^2}\right\|_\infty + \left\|\dfrac{\partial v}{\partial x}\right\|_\infty\right) + \Delta t\left\|\dfrac{\partial^2 v}{\partial t^2}\right\|_\infty\right], \\ 3N/4 \le i \le N - 1. \end{cases}$$

Arguing similarly for $(V_R - v)$, we obtain the following estimate for $i \in \{N/2 + 1, \ldots, N - 1\}$

$$|V_{R,i,k+1} - v(x_i, t_{k+1})| \le C\left(N^{-2} + \Delta t\right)(2 - x_i).$$

Now, we define the following two mesh functions on $S_x^N = \{x_i\}_0^N$:

$$S_i = \prod_{j=1}^{i}\left(1 + \frac{\alpha h_j}{\epsilon}\right), 1 \le i \le N/2, \quad \text{and}$$

$$Q_i = \prod_{j=1}^{N-i}\left(1 + \frac{\alpha h_j}{\epsilon}\right), \quad N/2 \le i \le N - 1$$

so that $S_0 = 1$ and $Q_N = 1$, where α is a positive constant.

Lemma 4.4 Let $\alpha \le \gamma/2$, then the functions S_i and Q_i satisfy

$$-L_\tau^{N,M}S_i \ge \begin{cases} \dfrac{C}{\varepsilon + \alpha H}S_i & \text{if } \ i = 1, \ldots, N\Big/4 \\ \dfrac{C}{\varepsilon + \alpha h}S_i & \text{if } \ i = N\Big/4 + 1, \ldots, N\Big/2 - 1 \end{cases}$$

and

$$-L_\tau^{N,M}Q_i \geq \begin{cases} \dfrac{C}{\varepsilon+\alpha h}Q_i & \text{if } i=N\Big/2+1,\ldots,3N\Big/4-1 \\ \dfrac{C}{\varepsilon+\alpha H}Q_i & \text{if } i=3N\Big/4,\ldots,N-1. \end{cases}$$

Proof. For $i = 1,\ldots,N/4$

$$-L_\tau^{N,M}S_i = -\frac{2\varepsilon}{\widehat{h}_i}\left[\left(\frac{S_{i+1}-S_i}{h_{i+1}}\right)-\left(\frac{S_i-S_{i-1}}{h_i}\right)\right] - a_{i-\frac{1}{2}}\left(\frac{S_i-S_{i-1}}{h_i}\right)+b_{i-\frac{1}{2}}S_{i-\frac{1}{2}}.$$

Since $S_i = \left(1+\frac{\alpha h_i}{\varepsilon}\right)S_{i-1}$, $S_i - S_{i-1} = \frac{\alpha h_i}{\varepsilon}S_{i-1}$ and $a_i \leq -\gamma_1 \leq -2\alpha$. Thus

$$-L_\tau^{N,M}S_i = -\frac{2\alpha}{\widehat{h}_i}(S_i-S_{i-1}) - a_{i-\frac{1}{2}}S_{i-1}+b_{i-\frac{1}{2}}S_{i-\frac{1}{2}} \geq \frac{C}{\varepsilon+\alpha h_i}S_i.$$

Next, for $i = N/4+1,\ldots,N/2-1$

$$-L_\tau^{N,M}S_i = -\varepsilon\delta_x^2 S_i - a_i D_x^0 S_i + b_i S_i + D_t^- S_i \geq \frac{C}{\varepsilon+\alpha h}S_i.$$

Now, for $i = N/2+1,\ldots,3N/4-1$

$$-L_\tau^{N,M}Q_i = -\varepsilon\delta_x^2 Q_i - a_i D_x^0 Q_i + b_i Q_i + D_t^- Q_i$$

$$= -\frac{\varepsilon}{h}\left[\left(\frac{Q_{i+1}-Q_i}{h_{i+1}}\right)-\left(\frac{Q_i-Q_{i-1}}{h_i}\right)\right] - a_i\left(\frac{Q_{i+1}-Q_{i-1}}{\widehat{h}_i}\right) + b_i Q_i.$$

Since $Q_{i+1}-Q_i = \frac{-\alpha h_{N-i}}{\varepsilon}Q_{i+1}$ and $a_i \geq \gamma_2 \geq 2\alpha$. Thus

$$-L_\tau^{N,M}Q_i \geq \frac{\alpha}{h}(Q_{i+1}-Q_i)+\frac{a_i}{2}\frac{\alpha}{\varepsilon}(Q_{i+1}-Q_{i-1}) \geq \frac{C}{(\varepsilon+\alpha h)}Q_i.$$

For $i = 3N/4,\ldots,N-1$

$$-L_\tau^{N,M}Q_i \geq -\frac{2\varepsilon}{\widehat{h}_i}\left[\left(\frac{Q_{i+1}-Q_i}{h_{i+1}}\right)-\left(\frac{Q_i-Q_{i-1}}{h_i}\right)\right] - a_{i+\frac{1}{2}}\left(\frac{Q_{i+1}-Q_i}{h_{i+1}}\right)$$

and $Q_{i+1}-Q_i = \frac{-\alpha h_{N-i}}{\varepsilon}Q_{i+1}$, $a_i \geq \gamma_2 \geq 2\alpha$. Therefore

$$-L_\tau^{N,M}Q_i \geq \frac{2\alpha}{\widehat{h}_i}(Q_{i+1}-Q_i)+a_{i+\frac{1}{2}}\frac{\alpha}{\varepsilon}Q_{i+1} \geq \frac{C}{(\varepsilon+\alpha h_{N-i})}Q_i.$$

Lemma 4.5 Let $\sigma_0 \geq 2/\alpha$. Then for $\{x_i\}_0^N$, the following inequalities are satisfied:

$$\left(\frac{S_i}{S_{N/2}}\right) \leq CN^{\frac{-4(1-2i)}{N}}, \quad i = N/4,\ldots,N/2-1 \tag{4.14}$$

and

$$\left(\frac{Q_i}{Q_{N/2}}\right) \leq CN^{-4\left(\frac{2i}{N-1}\right)}, \quad i = N/2+1,\ldots,3N/4. \tag{4.15}$$

Proof. Since $h_i = h$ for $i = N/4,\ldots,N/2-1$, thus

$$\left(\frac{S_i}{S_{N/2}}\right) = \left(1-\frac{\alpha h}{\varepsilon+\alpha h}\right)^{(N/2-i)}.$$

Taking log on both sides to obtain

$$\left(\frac{S_i}{S_{N/2}}\right) \leq \exp\left((N/2-i)\left(\frac{-\alpha h}{\varepsilon+\alpha h}\right)\right) \leq CN^{-4(1-2i/N)}.$$

Since sequence $N^{\frac{8(\alpha\sigma_0)^2(1-2i/N)(N^{-1}\ln N)}{(1+4\alpha\sigma_0 N^{-1}\ln N)}}$ is bounded. Further, for $i = N/2+1,\ldots,3N/4$

$$\left(\frac{Q_i}{Q_{N/2}}\right) = \frac{\prod_{j=1}^{N-i}\left(1+\frac{\alpha h}{\varepsilon}\right)}{\prod_{j=1}^{N/2}\left(1+\frac{\alpha h}{\varepsilon}\right)} = \left(1+\frac{\alpha h}{\varepsilon}\right)^{-(i-N/2)}.$$

Similarly, it is easy to follow that $\left(\frac{Q_i}{Q_{N/2}}\right) \leq CN^{-4(2i/N-1)}$.

Now, we will calculate the errors for the layers components W_L and W_R in $((0,1-\sigma]\cup[1+\sigma,2))\times(0,T]$.

Lemma 4.6 Let $\alpha \leq \gamma/2$ and $\sigma_0 \geq 2/\alpha$. Then under the assumptions (4.7) and (4.8), the errors associated to the layer components satisfy

$$\begin{cases} \left|W_{L,i,k+1}-w(x_i,t_{k+1})\right| \leq CN^{-2} & \text{for } i=1,\ldots,N/4 \\ \left|W_{R,i,k+1}-w(x_i,t_{k+1})\right| \leq CN^{-2} & \text{for } i=3N/4,\ldots,N-1. \end{cases}$$

Proof. From (4.13), we have

$$U_{N/2,k+1} = V_{R,N/2,k+1}+W_{R,N/2,k+1} = V_{L,N/2,k+1}+W_{L,N/2,k+1}$$

$$\left|U_{N/2,k+1}\right| \leq \left|V_{L,N/2,k+1}\right|+\left|W_{L,N/2,k+1}\right|.$$

Using Theorem 2.4 and Lemma 4.2 to obtain $\left|W_{N/2,k+1}\right| \leq C$. Consider

$$\phi_{L,i}^{k+1} = -C\left(\frac{S_i}{S_{N/2}}\right) \quad \text{for} \quad i=0,\ldots,N/2.$$

For $i = 0,\ldots,N/2$ and using Lemma 4.4, we have

$$L_\tau^{N,M}\left(\phi_{L,i}^{k+1} \pm W_{L,i,k+1}\right) = L_\tau^{N,M}\phi_{L,i}^{k+1} \pm L_\tau^{N,M}W_{L,i,k+1} \geq \frac{C}{\varepsilon+\alpha H}\frac{S_i}{S_{N/2}} \geq 0$$

and $W_{L,0,k+1} = 0 = W_{L,i,0}$ for $i \leq N/2$. Thus

$$\phi_{L,0}^{k+1} \pm W_{L,0,k+1} = \phi_{L,0}^{k+1} = \frac{-CS_0}{S_{N/2}} \leq 0.$$

Similarly, we can show $\phi_{L,i}^0 \pm W_{L,i,0} \leq 0$. Therefore using Lemma 4.1, we get

$$|W_{L,i,k+1}| \leq C\left(\frac{S_i}{S_{N/2}}\right) \quad \text{for} \quad i = 1,\ldots,N/2-1. \tag{4.16}$$

Further

$$|W_{L,i,k+1}| \leq C\left(\frac{S_i}{S_{N/2}}\right) \leq C\left(\frac{S_{N/4}}{S_{N/2}}\right)$$

and from Lemma 4.5, we have

$$|W_{L,i,k+1}| \leq CN^{-2}. \tag{4.17}$$

Since $\alpha \leq \gamma/2 \leq \gamma$, $\sigma = \sigma_0 \varepsilon \ln N$, $\sigma_0 \geq 2/\alpha$ and using Theorem 2.4, we have

$$|w(x_i,t_{k+1})| \leq C\exp\left(\frac{-(1-x_i)\gamma_1}{\varepsilon}\right) \leq CN^{-2} \text{ for } i=1,\ldots,N/4. \tag{4.18}$$

Combine (4.17) and (4.18) to obtain

$$|W_{L,i,k+1} - w(x_i,t_{k+1})| \leq CN^{-2} \quad \text{for} \quad i=1,\ldots,N/4.$$

Now, consider $\phi_{R,i}^{k+1} = -C\left(\frac{Q_i}{Q_{N/2}}\right)$ for $i = N/2,\ldots,N$. Then using Lemma 4.4, for $i = 3N/4,\ldots,N-1$ to obtain

$$L_\tau^{N,M}\left(\phi_{R,i}^{k+1} \pm W_{R,i,k+1}\right) = L_\tau^{N,M}\phi_{R,i}^{k+1} \pm L_\tau^{N,M}W_{R,i,k+1} \geq \frac{C}{\varepsilon+\alpha H}\frac{Q_i}{Q_{N/2}} \geq 0$$

and $W_{R,N,k+1} = 0 = W_{R,i,0}$. Thus $\phi_{R,N}^{k+1} \pm W_{R,N,k+1} = \phi_{R,N}^{k+1} \leq 0$.

Similarly, we can show $\phi_{R,i}^0 \pm W_{R,i,0} \leq 0$. Thus using Lemma 4.1, we have $\phi_{R,i}^{k+1} \pm W_{R,i,k+1} \leq 0$ for $i = N/2,\ldots,N$. Thus

$$|W_{R,i,k+1}| \leq \left|\phi_{R,i}^{k+1}\right| \leq C\left(\frac{Q_i}{Q_{N/2}}\right) \text{ for } i=N/2+1,\ldots,N-1 \leq C\left(\frac{Q_{N/4}}{Q_{N/2}}\right) \text{ for } i=3N/4,\ldots,N-1$$

since Q is decreasing and from Lemma 4.5, we have $\left(\frac{Q_{N/4}}{Q_{N/2}}\right) \leq CN^{-2}$. Thus

$$|W_{R,i,k+1}| \leq CN^{-2}. \tag{4.19}$$

Now, using Theorem 2.4 and doing the same calculation as we do for W_L, we have

$$|w(x_i,t_{k+1})| \leq C\exp\left(\frac{-(1-x_i)\gamma_2}{\varepsilon}\right) \leq CN^{-2}. \tag{4.20}$$

Using (4.19) and (4.20), we obtain $\left|W_{R,i,k+1} - w(x_i, t_{k+1})\right| \leq CN^{-2}$.

Next, we will state and prove some lemmas required to obtain ε uniform error bounds.

Lemma 4.7 The following inequalities hold true:

$$\exp(-\alpha(1-x_i)/\varepsilon) \leq \left(\frac{S_i}{S_{N/2}}\right), \quad i=1,\ldots,N/2-1 \tag{4.21}$$

and

$$\exp(-\alpha(x_i-1)/\varepsilon) \leq \left(\frac{Q_i}{Q_{N/2}}\right), \quad i=N/2+1,\ldots,N-1. \tag{4.22}$$

Proof. Follows from Lemma 2.5 in [40], for each j, we obtain

$$\exp\left(\frac{-\alpha h_j}{\varepsilon}\right) = \left(\exp\left(\frac{\alpha h_j}{\varepsilon}\right)\right)^{-1} \leq \left(1+\frac{\alpha h_j}{\varepsilon}\right)^{-1}. \tag{4.23}$$

Multiplying (4.23), for $j = i+1,\ldots,N/2$ to obtain

$$\prod_{i+1}^{N/2}\exp\left(\frac{-\alpha h_j}{\varepsilon}\right) \leq \prod_{i+1}^{N/2}\left(1+\frac{\alpha h_j}{\varepsilon}\right)^{-1} \leq \frac{S_i}{S_{N/2}}.$$

Further, multiplying (4.23), for $j = N-(i+1),\ldots, N/2$, we have

$$\prod_{N-(i+1)}^{N/2}\exp\left(\frac{-\alpha h_j}{\varepsilon}\right) \leq \prod_{N-(i+1)}^{N/2}\left(1+\frac{\alpha h_j}{\varepsilon}\right)^{-1} \leq \frac{Q_i}{Q_{N/2}}.$$

Lemma 4.8 The difference operators D_x^B and D_x^F satisfy the following inequalities:

$$D_x^B S_{N/2} \geq \frac{C}{\varepsilon+\alpha h} S_{N/2}, \quad -D_x^F Q_{N/2} \geq \frac{C}{\varepsilon+\alpha h} Q_{N/2}.$$

Proof. Since $(S_i - S_{i-1}) = \frac{\alpha h_i}{\varepsilon} S_{i-1}$, we have

$$D_x^B S_{N/2} = \frac{1}{2h}\left(S_{N/2-2} - 4S_{N/2-1} + 3S_{N/2}\right)$$
$$\geq \frac{\alpha}{\varepsilon}\frac{S_{N/2}}{\left(1+\frac{\alpha h}{\varepsilon}\right)^2}\left(1+\frac{\alpha h}{\varepsilon}\right) \geq \frac{C}{\varepsilon+\alpha h} S_{N/2}.$$

Since $Q_{i+1} - Q_i = \frac{-\alpha h_{N-i}}{\varepsilon} Q_{i+1}$, we have

$$-D_x^F Q_{N/2} = \frac{1}{2h}\left(Q_{N/2+2} - 4Q_{N/2+1} + 3Q_{N/2}\right)$$
$$\geq \frac{\alpha}{\varepsilon}\frac{\left(1+\frac{\alpha h}{\varepsilon}\right)}{\left(1+\frac{\alpha h}{\varepsilon}\right)^2} Q_{N/2} \geq \frac{C}{\varepsilon+\alpha h} Q_{N/2}.$$

Theorem 4.9 Let u and $U_{i,k+1}$ be the solution of (2.1) and (4.5), respectively. Then under the assumptions (4.7), (4.8) and $\alpha \leq \gamma/2$ and $\sigma_0 \geq 2/\alpha$, we get

$$|U_{i,k+1} - u(x_i, t_{k+1})| \leq \begin{cases} C\left(N^{-2} + \Delta t\right) & \text{for } i = 1, \ldots, N/4, 3N/4, \ldots, N-1 \\ C\left(N^{-2}\ln^2 N + \Delta t\right) & \text{for } i = N/4+1, \ldots, 3N/4-1. \end{cases} \tag{4.24}$$

Proof. We compute the error separately in the layer region and outside the layer region.

Case 1: For $i = 1, \ldots, N/4, 3N/4, \ldots, N-1$.Follows from the triangle inequality to $U - u$ and using Lemma 4.3 and Lemma 4.6 to obtain

$$|U_{i,k+1} - u(x_i, t_{k+1})| \leq |V_{L,i,k+1} - v(x_i, t_{k+1})|$$
$$+ |W_{L,i,k+1} - w(x_i, t_{k+1})|$$
$$\leq C\left(N^{-2} + \Delta t\right).$$

Case 2: Here, we need to find out the error estimate $\left|U_{i,k+1} - u(x_i, t_{k+1})\right|$ for $i = N/4+1, \ldots, 3N/4-1$. For $i = N/4, 3N/4$, From Case 1

$$|U_{i,k+1} - u(x_i, t_{k+1})| \leq C\left(N^{-2} + \Delta t\right). \tag{4.25}$$

Consider

$$L_\tau^{N,M}\left(U_{i,k+1} - u(x_i, t_{k+1})\right) = L_\tau^{N,M} U_{i,k+1} - L_\tau^{N,M} u(x_i, t_{k+1})$$
$$= \left(\varepsilon\left(\frac{\partial^2}{\partial x^2} - \delta x^2\right) + a\left(\frac{\partial}{\partial x} - D_x^0\right) - \left(\frac{\partial}{\partial t} - D_t^-\right)\right) u(x_i, t_{k+1}).$$

Using Theorem 2.4, we have

$$\left|L_\tau^{N,M}(U_{i,k+1} - u(x_i, t_{k+1}))\right| \leq h \int_{x_{i-1}}^{x_{i+1}} \left(\varepsilon\left|\frac{\partial^4 u}{\partial x^4}\right| + \left|\frac{\partial^3 u}{\partial x^3}\right|\right) dx$$
$$+ C\Delta t \left\|\frac{\partial^2 u}{\partial t^2}\right\|_\infty$$
$$\leq C\left[h^2 + \frac{h}{\varepsilon^2}\exp\left(\frac{-(1-x_i)\gamma_1}{\varepsilon}\right)\sinh\left(\frac{h\gamma_1}{\varepsilon}\right)\right] + \Delta t. \tag{4.26}$$

From (3.5), we have

$$2\sigma_0\gamma^* \leq \frac{N}{\ln N} 4\sigma_0\gamma_1 \leq \frac{2N}{\ln N}, \quad \frac{4\sigma\gamma_1}{\varepsilon \ln N} \leq \frac{2N}{\ln N} \text{ and } \frac{h\gamma_1}{\varepsilon} \leq 2.$$

Since $\sinh x \leq cx$, $x \in [0, 2]$. This implies $\sinh\left(\frac{h\gamma_1}{\varepsilon}\right) \leq C\left(\frac{\gamma_1 h}{\varepsilon}\right)$. Therefore (4.26) becomes

$$\left|L_\tau^{N,M}\left(U_{i,k+1} - u(x_i, t_{k+1})\right)\right|$$
$$\leq C\left[\left(h^2 + \frac{h^2}{\varepsilon^3}\exp\left(\frac{-(1-x_i)\gamma_1}{\varepsilon}\right)\right) + \Delta t\right]. \tag{4.27}$$

Similarly, for $i = N/2+1, \ldots, 3N/4-1$

$$\left|L_\tau^{N,M}\left(U_{i,k+1} - u(x_i, t_{k+1})\right)\right|$$
$$\leq C\left[\left(h^2 + \frac{h^2}{\varepsilon^3}\exp\left(\frac{-(x_i-1)\gamma_2}{\varepsilon}\right)\right) + \Delta t\right]. \tag{4.28}$$

Further at $x_{N/2} = 1$

$$\left|L_\tau^{N,M}\left(U_{N/2,k+1} - u\left(x_{N/2}, t_{k+1}\right)\right)\right|$$
$$= \left|\tilde{f}_{\tau,N/2,k+1} - L_\tau^{N,M} u_{N/2,k+1}\right| \leq C\left(\frac{h^2}{\varepsilon^3} + \Delta t\right).$$

Define the following discrete function

$$\Theta_i^k = \begin{cases} -C\left(N^{-2}+\Delta t\right)(1+(x_i-(1-\sigma)))-C\dfrac{h^2}{\varepsilon^2}\left(\dfrac{S_i}{S_{N/2}}\right) & \text{for} \quad i=N/4,\ldots,N/2 \\ -C\left(N^{-2}+\Delta t\right)(1+((1+\sigma)-x_i))-C\dfrac{h^2}{\varepsilon^2}\left(\dfrac{Q_i}{Q_{N/2}}\right) & \text{for} \quad i=N/2+1,\ldots,3N/4. \end{cases}$$

Then

$$L_\tau^{N,M}\Theta_i^{k+1} = \begin{cases} -C\left(N^{-2}+\Delta t\right)L_\tau^{N,M}(x_i)-C\dfrac{h^2}{\varepsilon^2}\left(\dfrac{L_\tau^{N,M}S_i}{S_{N/2}}\right) & \text{for} \quad i=N/4,\ldots,N/2 \\ C\left(N^{-2}+\Delta t\right)L_\tau^{N,M}(x_i)-C\dfrac{h^2}{\varepsilon^2}\left(\dfrac{L_\tau^{N,M}Q_i}{Q_{N/2}}\right) & \text{for} \quad i=N/2+1,\ldots,3N/4 \end{cases}$$

$$= \begin{cases} -C\left(N^{-2}+\Delta t\right)(a_i-b_ix_i)-C\dfrac{h^2}{\varepsilon^2}\left(\dfrac{L_\tau^{N,M}S_i}{S_{N/2}}\right) & \text{for} \quad i=N/4,\ldots,N/2 \\ C\left(N^{-2}+\Delta t\right)(a_i-b_ix_i)-C\dfrac{h^2}{\varepsilon^2}\left(\dfrac{L_\tau^{N,M}Q_i}{Q_{N/2}}\right) & \text{for} \quad i=N/2+1,\ldots,3N/4. \end{cases}$$

Using assumption $\alpha \le \gamma/2$, Lemma 4.6 and Lemma 4.7 to obtain

$$L_\tau^{N,M}\Theta_i^{k+1} \ge \begin{cases} C\gamma_1\left(N^{-2}+\Delta t\right)+C\dfrac{h^2}{\varepsilon^2}\exp(-(1-x_i)\gamma_1/\varepsilon) & \text{for} \quad i=N/4,\ldots,N/2 \\ C\gamma_2\left(N^{-2}+\Delta t\right)+C\dfrac{h^2}{\varepsilon^2}\exp(-(x_i-1)\gamma_2/\varepsilon) & \text{for} \quad i=N/2+1,\ldots,3N/4. \end{cases}$$

Assumption (4.7) implies $\frac{h}{\varepsilon} \le \frac{2}{\gamma^*}$ and using Lemma 4.8, we get

$$L_\tau^{N,M}\Theta_i^{k+1} \ge \left(D_x^F - D_x^B\right)\Theta_i^{k+1} \ge 2C\left(N^{-2}+\Delta t\right)+2C\frac{h^2}{\varepsilon^3}. \tag{4.29}$$

Therefore, it follows from (4.25)−(4.29) that

$$\begin{cases} L_\tau^{N,M}\Theta_i^{k+1} \geq \left|L_\tau^{N,M}(U_{i,k+1}-u(x_i,t_{k+1}))\right| & \text{for } i=N/4+1,\ldots,3N/4-1 \\ -\Theta_i^{k+1} \geq |U_{i,k+1}-u(x_i,t_{k+1})| & \text{for } i=N/4,3N/4 \text{ and} \\ -\Theta_i^{0} \geq |U_{i,0}-u(x_i,t_0)|. \end{cases}$$

Then applying discrete maximum principle to $\Theta_i^{k+1} \pm (U_{i,k+1}-u(x_i,t_{k+1}))$ over the domain $\overline{D}^{N,M} \cap ([1-\sigma,1+\sigma]\times[0,T])$ we obtain for $i = N/4+1, \ldots, 3N/4-1$

$$|U_{i,k+1}-u(x_i,t_{k+1})| \leq C\left(\frac{h^2}{\varepsilon^2}+\Delta t\right) \leq C\left(N^{-2}\ln^2 N+\Delta t\right).$$

5. Numerical illustration

The performance of the proposed method is examined in this section and the theoretical estimates are numerically verified. We consider two test problems for numerical computations.

Example 5.1 Consider the following singularly perturbed problem of class $(2.1)-(2.3)$

$$\begin{cases} \varepsilon u_{xx}+a(x)u_x-x(2-x)u(x,t)-u_t=f(x,t)+u(x-1,t), & (x,t)\in(0,2)\times(0,2], \\ u(x,0)=0, \quad x\in[0,2], \\ u(x,t)=t^2, \quad (x,t)\in[-1,0]\times[0,2], \\ u(2,t)=0, \quad t\in(0,2], \end{cases}$$

where

$$a(x)=\begin{cases} -(2+x(2-x)), & x\in[0,1] \\ (2+x(2-x)), & x\in(1,2] \end{cases}$$

and

$$f(x,t)=\begin{cases} 2(1+x^2)t^2, & (x,t)\in[0,1]\times[0,2] \\ 3(1+x^2)t^2, & (x,t)\in(1,2]\times[0,2] \end{cases}$$

Example 5.2 Consider the following singularly perturbed problem of class $(2.1)-(2.3)$

$$\begin{cases} \varepsilon u_{xx}+a(x)u_x-5u(x,t)-u_t=f(x,t)+2u(x-1,t), & (x,t)\in(0,2)\times(0,2], \\ u(x,0)=0, \quad x\in[0,2], \\ u(x,t)=0, \quad (x,t)\in[-1,0]\times[0,2], \\ u(2,t)=0, \quad t\in(0,2], \end{cases}$$

where

$$a(x)=\begin{cases} -(4+x^2), & x\in[0,1] \\ (6-x^2), & x\in(1,2] \end{cases}$$

and

$$f(x,t)=\begin{cases} 4xt^2\exp(-t), & (x,t)\in[0,1]\times[0,2] \\ 4(2-x)t^2\exp(-t), & (x,t)\in(1,2]\times[0,2] \end{cases}$$

The exact solutions for the problems are unknown for comparison. Therefore, we use the double mesh principle to estimate the error. The maximum point−wise error ($E_\varepsilon^{N,\triangle t}$) and order of convergence ($R_\varepsilon^{N,\triangle t}$) are calculated using:

$$E_\varepsilon^{N,\triangle t} := \max\left|U^{N,\triangle t}(x_i,t_{k+1})-\tilde{U}^{2N,\triangle t/2}(x_i,t_{k+1})\right|,$$
$$R_\varepsilon^{N,\triangle t} := \log_2\left(\frac{E_\varepsilon^{N,\triangle t}}{E_\varepsilon^{2N,\triangle t/2}}\right),$$

where $U^{N,\triangle t}(x_i,t_{k+1})$ and $\tilde{U}^{2N,\triangle t/2}(x_i,t_k)$ are the approximate solutions obtained on the mesh $D^{N,M}$ and $D^{2N,2M}$, respectively. When, the perturbation parameter approaches zero, the problem's solution exhibits turning point behaviour (Figs. 1−4).

Maximum absolute error and order of convergence for problems 5.1 and 5.2 are tabulated in Tables 1 and 2. Moreover, the maximum absolute errors for problems 5.1 and 5.2 are plotted in Figs. 5 and 6, respectively. The surface plot of the numerical solution for problems 5.1 and 5.2 are plotted in Figs. 1 and 3, respectively. Also, the numerical solutions at final time step ($t = 2$) for different values of ε are displayed in Figs. 2 and 4.

The numerical results tabulated in Tables 1 and 2 do not clearly depict the theoretical order of convergence for spatial discretization. It is to be noted that the error in numerical solution is due to spatial and temporal discretization. As a consequence, the errors given in Tables 1 and 2 are a combination of temporal and spatial errors, with the layer regions playing a significant role.

The hybrid difference scheme improves accuracy in space only. To verify this, we performed numerical experiments for $M = N^2$ and numerical results are tabulated in Table 3. In Table 4, we have fixed $\varepsilon = 2^{-6}$ and $N = 512$ and reduce Δt by half and the associated errors are presented at different values of x. It can be observed that the errors reduce by almost half which confirms the first-order convergence in time. Figures 7 and 8 have also been illustrated to show the errors in layer region and outside layer region. This demonstrates that the numerical method is second-order spatially accurate outside of the interior layer and the errors are reduced in the layer region as claimed in Theorem 4.9. The implementation of the proposed method is done in MATLAB R2015b (The Mathworks, Inc.). The program code is uploaded in GitHub and the URL of the source code is http://shorturl.at/gjlCS.

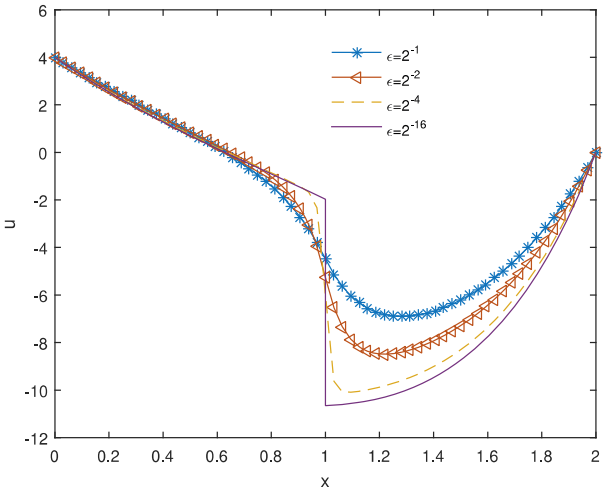


Fig. 2. Numerical solution of Example 5.1 at $t = 2$ for different values of ε when $N = 128$.

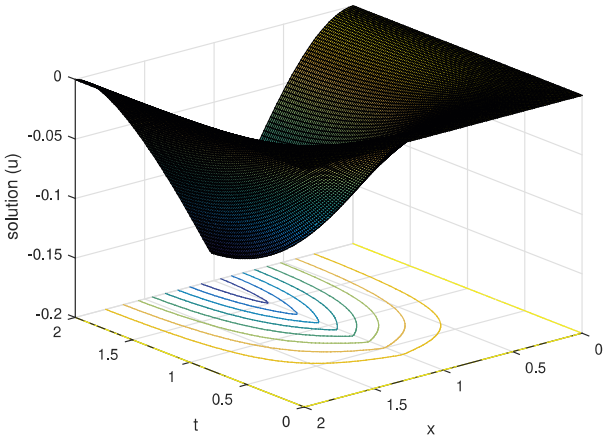


Fig. 3. Numerical solution of Example 5.2 for $\varepsilon = 2^{-4}$ when $M = N = 128$.

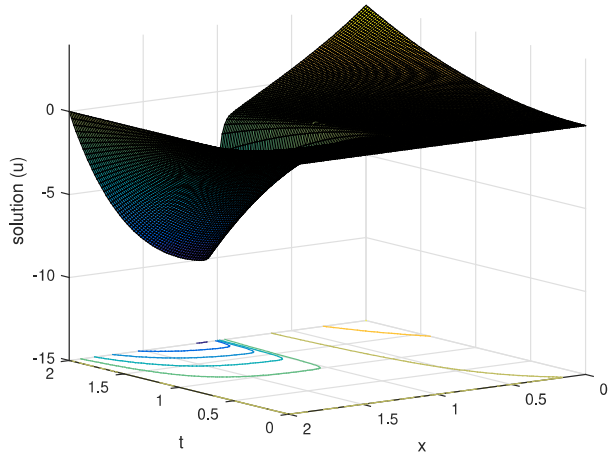


Fig. 1. Numerical solution of Example 5.1 for $\varepsilon = 2^{-4}$ when $M = N = 128$.

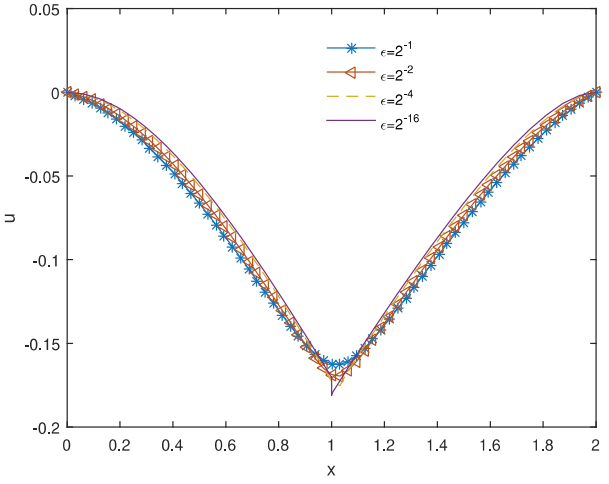


Fig. 4. Numerical solution of Example 5.2 at $t = 2$ for different values of ε when $N = 128$.

Table 1. Maximum absolute error and order of convergence for Example 5.1 for different values of ε, M and N when M = N.

N	$\varepsilon = 2^{-2}$	2^{-4}	2^{-6}	2^{-8}	2^{-10}	2^{-12}
32	1.099e-01	8.697e-01	8.723e-01	8.712e-01	8.709e-01	8.708e-01
	1.6274	1.5675	1.2466	1.2572	1.2601	1.2609
64	3.558e-02	2.934e-01	3.676e-01	3.644e-01	3.635e-01	3.633e-01
	1.3320	2.0478	1.5866	1.6018	1.6060	1.6071
128	1.413e-02	7.096e-02	1.223e-01	1.200e-01	1.194e-01	1.192e-01
	1.1428	1.9332	1.7309	1.7302	1.7308	1.7311
256	6.400e-03	1.858e-02	3.687e-02	3.619e-02	3.598e-02	3.592e-02
	1.0336	1.7107	1.6093	1.6294	1.6359	1.6376
512	3.126e-03	5.676e-03	1.208e-02	1.169e-02	1.157e-02	1.154e-02
	1.0066	1.5189	1.5926	1.6132	1.6207	1.6241

Table 2. Maximum absolute error and order of convergence for Example 5.2 for different values of ε, M and N when M = N.

N	$\varepsilon = 2^{-2}$	2^{-4}	2^{-6}	2^{-8}	2^{-10}	2^{-12}
32	1.886e-03	4.150e-03	2.485e-03	2.146e-03	2.288e-03	2.323e-03
	1.4071	1.6239	1.1244	1.0741	1.0986	1.1044
64	7.113e-04	1.346e-03	1.139e-03	1.019e-03	1.068e-03	1.080e-03
	1.5199	1.6981	1.5245	1.3442	1.3740	1.3811
128	2.480e-04	4.149e-04	3.961e-04	4.015e-04	4.122e-04	4.149e-04
	1.5620	1.4068	1.5615	1.6111	1.6576	1.6714
256	8.400e-05	1.565e-04	1.342e-04	1.313e-04	1.304e-04	1.302e-04
	1.5102	1.5403	1.5137	1.6794	1.7007	1.7062
512	2.949e-05	5.380e-05	4.700e-05	4.101e-05	4.014e-05	3.991e-05
	1.4117	1.5972	1.4798	1.4726	1.4368	1.4272

6. Concluding remark

Singularly perturbed parabolic functional differential equation with discontinuous coefficient and source term is numerically solved. The problem's solution takes into account not just the present state of the physical system but also its history. The simultaneous presence of discontinuous data and delay makes the problem stiff. In the limiting case, the solution of the problem exhibits multi-scale character. There are narrow regions where solution derivatives grow exponentially and exhibit turning point behaviour, leading to sharp interior layers across discontinuities.

A hybrid numerical scheme composed of a central difference scheme in the layer region and a midpoint upwind scheme outside the layer region is used to discretize space variable over a specially generated mesh. Whereas an implicit finite difference scheme is used to discretize the time variable. The mesh has been chosen so that most of the mesh

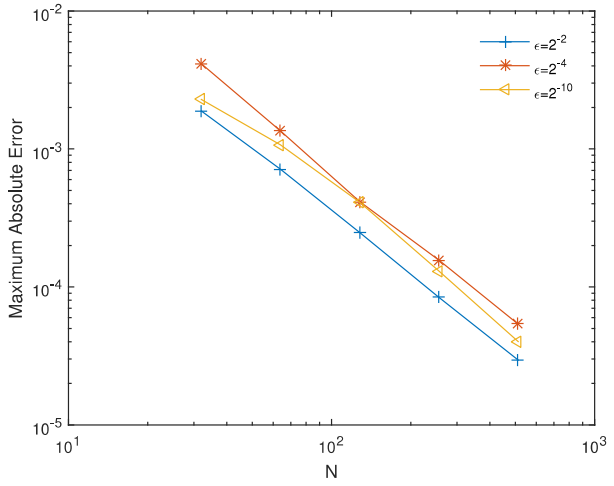


Fig. 5. Error plot for Example 5.1.

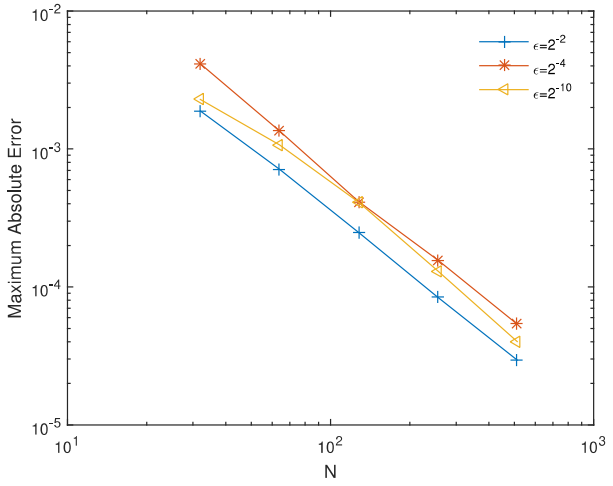


Fig. 6. Error plot for Example 5.2.

Table 3. Maximum absolute error and order of convergence for Example 5.1 and 5.2 for different values of M and N when $M = N^2$ and $\varepsilon = 2^{-10}$.

For Example 5.1				For Example 5.2		
N	*left region* $[0, 1-\sigma]$	*interior layer region* $(1-\sigma, 1+\sigma)$	*righ tregion* $[1+\sigma, 2]$	*left region* $[0, 1-\sigma]$	*interior layer region* $(1-\sigma, 1+\sigma)$	*right region* $[1+\sigma, 2]$
32	7.602e-03	9.073e-01	2.237e-02	1.863e-04	2.215e-03	6.044e-05
	1.9665	1.3111	1.9344	1.9878	1.1088	1.8829
64	1.973e-03	3.656e-01	5.687e-03	4.652e-05	1.026e-03	1.632e-05
	1.9760	1.6430	1.9033	1.9624	1.3972	1.7645
128	5.083e-04	1.170e-01	1.478e-03	1.182e-04	3.895e-04	4.790e-06
	1.9592	1.7816	1.8513	1.9592	1.7332	1.5841
256	1.323e-04	3.404e-02	3.992e-04	3.064e-06	1.186e-04	1.594e-06
	1.9438	1.9236	1.8205	1.9382	1.8560	1.5632

Table 4. Maximum absolute error and order of convergence for Example 5.2 for different values of M and x when $N = 512$ and $\varepsilon = 2^{-6}$.

x	$M = 32$	64	128	256	512
$x_{N/2+1}$	3.837e-04	2.058e-04	1.071e-04	5.486e-05	2.773e-05
	0.8988	0.9423	0.9651	0.9843	0.9981
$x_{N/2+4}$	3.999e-04	2.122e-4	1.093e-04	6.530e-05	3.225e-05
	0.9142	0.9571	0.7432	1.0178	1.0184

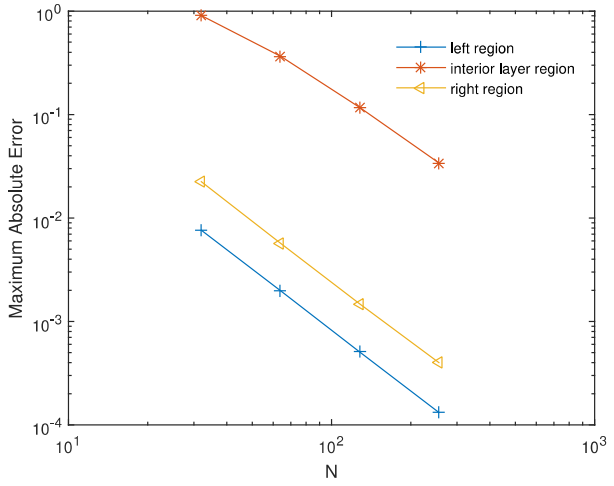


Fig. 7. Error plot of the spatial order of convergence for example 5.1.

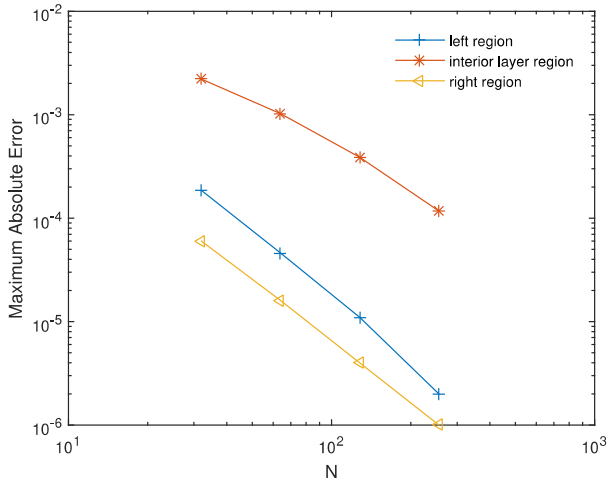


Fig. 8. Error plot of the spatial order of convergence for example 5.2.

points remain in the regions with rapid transitions. The proposed numerical method has been analyzed for consistency, stability and convergence. Extensive theoretical analysis is performed to obtain consistency and error estimates. It is found that the method proposed is unconditionally stable, and the convergence obtained is parameter uniform. Numerical illustrations are presented for two test examples that demonstrate the effectiveness of the technique. Convergence obtained in practical satisfies theoretical predictions.

References

[1] der Heiden U, Walther H-O. Existence of chaos in control systems with delayed feedback. J Differ Equ 1983;47:273−95.
[2] Ansari AR, Bakr SA, Shishkin GI. A parameter robust finite difference method for singularly perturbed delay parabolic partial differential equations. J Comput Appl Math 2007;205: 552−66.
[3] Avudai-Selvi P, Ramanujam N. A parameter uniform difference scheme for singularly perturbed parabolic delay differential equation with Robin type boundary condition. Appl Math Comput 2017;296:101−15.
[4] Barbu L, Morosanu G. Singularly perturbed boundary value problems. Basel: Birkh a usar; 2007.
[5] Bashier EBM, Patidar KC. A fitted numerical method for a system of partial delay differential equations. Comput Math Appl 2011;61:1475−92.
[6] Bashier EBM, Patidar KC. A novel fitter operator finite difference method for a singularly perturbed delay parabolic partial differential equation. Appl Math Comput 2011;217: 4728−39.
[7] Bobisud L. Second-order linear parabolic equations with a small parameter. Arch Ration Mech Anal 1967;27:385−97.
[8] Clavero C, Gracia JL. A high order HODIE finite difference scheme for 1D parabolic singularly perturbed reaction diffusion problems. Appl Math Comput 2012;218:5067−80.
[9] Das A, Natesan S. Uniformly convergent hybrid numerical scheme for singularly perturbed delay parabolic convection

diffusion problems on Shishkin mesh. Appl Math Comput 2015;271:168−86.
[10] Erneux T. Applied delay differential equations. Surv. Tutor. Appl. Math. Sci. 2009;3. Springer, New York.
[11] Farrell PA, Hegarty AF, Miller JJH, O'Riordan E, Shishkin GI. Robust computational Techniques for boundary layers. Boca Raton: Chapman and Hall/CRC; 2000.
[12] Franz AL, Roy R, Shaw LB, Schwartz IB. Effect of multiple time delays on intensity fluctuation dynamics in fiber ring lasers. Phys Rev E 2008;78:16208.
[13] Gowrisankar S, Natesan S. A robust numerical scheme for singularly per-turbed delay parabolic initial boundary value problems on equidistributed grids. Electron Trans Numer Anal 2014;41:376−95.
[14] Gowrisankar S, Natesan S. ε-uniformly convergent numerical scheme for singularly perturbed delay parabolic partial differential equations. Int J Comput Math 2017;94:902−21.
[15] Gupta V, Kumar M, Kumar S. Higher order numerical approximation for time dependent singularly perturbed differential-difference convection-diffusion equations. Numer Methods Part Differ Equ 2018;34:357−80.
[16] Gurney WSC, Blythe SP, Nisbet RM. Nicholson's blowies revisited. Nature 1980;287:17−21.
[17] Hale JK, Lunel SMV. Introduction to functional differential equations. Berlin: Springer; 1993.
[18] Hemker PW, Shishkin GI, Shishkina LP. High-order time-accurate schemes for singularly perturbed parabolic convection-diffusion problems with robin boundary conditions. Comput Methods Appl Math 2002;2:3−25.
[19] Ikeda K. Multiple-valued stationary state and its stability of the transmitted light by a ring cavity system. Opt Commun 1979;30:257−61.
[20] Kaushik A, Sharma KK, Sharma M. A parameter uniform difference scheme for parabolic partial differential equation with a retarded argument. Appl Math Model 2010;34: 4232−42.
[21] Kaushik A and Sharma N. An adaptive difference scheme for parabolic delay differential equation with discontinuous coefficients and interior layers, J Differ Equ Appl 26 (2020), 11-12.
[22] Kumar K, Chakravarthy PP, Ramos H, Vigo-Aguiar J. A stable finite difference scheme and error estimates for parabolic singularly perturbed PDEs with shift parameters. J Comput Appl Math 2022;405:15. Paper No. 113050.
[23] Kumar M, Kumar S. High order robust approximations for singularly perturbed semilinear systems. Appl Math Model 2012;36:3570−9.
[24] Kumar M, Rao SCS. High order parameter robust numerical method for singularly perturbed reaction diffusion problems. Appl Math Comput 2010;216:1036−46.
[25] Kumar S, Kumar M. High order parameter uniform discretization for singularly perturbed parabolic partial differential equations with time delay. Comput Math Appl 2014;68:1355−67.
[26] Ladyzhenskaya OA, Solonnikov VA, Ural'ceva NN. Linear and quasilinear equations of parabolic type. Translations of mathematical monographs, 23. Providence, R.I.: American Mathematical Society; 1968. Translated from the Russian by S. Smith.
[27] Larger L, Penkovsky B, Maistrenko Y. Virtual chimera states for delayed- feedback systems. Phys Rev Lett 2013;111:54103.
[28] Linß T. Layer adapted meshes for reaction convection diffusion problems. Berlin: Springer; 2010.
[29] Mackey MC, Glass L. Oscillation and chaos in physiological control systems. Sci New Ser 1977;197:287−9.
[30] Marconi M, Javaloyes J, Barland S, Balle S, Giudici M. Vectorial dissipative solitons in vertical-cavity surface-emitting lasers with delays. Nat Photonics 2015;9:450−5.
[31] Marcus CM, Westervelt M. Stability of analog neural networks with delay. Phys Rev 1989;39:347−59.
[32] Miller JJH, O'Riordan E, Shishkin GI. Fitted numerical Methods for singular perturbation problems. Singapore: World Scientific; 2012.
[33] Miller JJH, O 'Riordan E, Shishkin GI, Shishkina LP. Fitted mesh methods for problems with parabolic boundary layers. Math Proc Roy Ir Acad 1998;98:173−90.
[34] Ramesh VP and Priyanga B. Higher order uniformly convergent numerical algorithm for time-dependent singularly perturbed differential-difference equations, Differ Equ Dyn Syst 29 (2021), 239-263.
[35] Roos HG, Stynes M, Tobiska L. Numerical methods for singularly perturbed differential equations: convection-diffusion and flow problems. New york: Springer- Verlag; 1996.
[36] Roos HG, Stynes M, Tobiska L. Robust numerical methods for singularly perturbed differential equations. Berlin: Springer; 2008.
[37] Scholl E, Schuster H. Handbook of chaos control. Weinheim: Wiley-VCH; 2008.
[38] Shivhare M, Chakravarthy PP, Ramos H, Vigo-Aguiar J. Quadratic b- spline collocation method for time dependent singularly perturbed differential- difference equation arising in the modeling of neuronal activity. Numer Methods Part Differ Equ 2021:1−22.
[39] Soriano MC, Garcia-Ojalvo J, Mirasso C, Fischer I. Complex photonics: Dynamics and applications of delay-coupled semiconductors lasers. Rev Mod Phys 2013;85:421−70.
[40] Stynes M, Roos HG. The midpoint upwind scheme. Appl Numer Math 1997;23:361−74.
[41] Stynes M, Tobiska L. The SDFEM for a convection diffusion problem with a boundary layer: Optimal error analysis and enhancement of accuracy. SIAM J Numer Anal 2003;41: 1620−42.
[42] Wu J. Theory and applications of partial functional differential equations. Berlin: Springer; 1996.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/362469247

A Performance Perspective of Live Migration of Virtual Machine in Cloud Data Center with Future Directions

Article *in* International Journal of Wireless and Microwave Technologies · January 2022

DOI: 10.5815/ijwmt.2022.04.04

CITATIONS
0

READS
6

3 authors, including:

Divya Kapil
GEHU
14 PUBLICATIONS **201** CITATIONS
SEE PROFILE

Vishan Kumar Gupta
Graphic Era Deemed to be University
13 PUBLICATIONS **104** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

A Performance Perspective of Live Migration of Virtual Machine in Cloud Data Center with Future Directions View project

All content following this page was uploaded by Vishan Kumar Gupta on 04 August 2022.

The user has requested enhancement of the downloaded file.

I.J. Wireless and Microwave Technologies, **2022, 4, 48-56**
Published Online August 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijwmt.2022.04.04

A Performance Perspective of Live Migration of Virtual Machine in Cloud Data Center with Future Directions

Divya Kapil[1], Saurabh Kumar Mishra[2]
[1]School of Computing, Graphic Era Hill University, Dehradun, India
[2]Dept. of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

Vishan Kumar Gupta
Dept. of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India
Email: vishangupta@gmail.com

Received: 26 March 2022; Accepted: 26 May 2022; Published: 08 August 2022

Abstract: Live virtual machine migration is a valuable feature for the virtualized data center or cloud environment. This is the process to migrate running virtual machines from one physical host to another host. Live virtual machine migration can be used to provide various benefits such as server consolidation, energy-saving, and maintenance. It is a valuable feature for the virtualized data center or cloud environment. Cloud computing provides IT capabilities as a service and its key technology is virtualization. The key benefit of virtualization is to offer better resource employment by executing various VMs on the same physical system. In this research, we analyze the performance of the various hypervisors based on their migration features and compares the migration feature. Hypervisors are Xen, VMWare, KVM, and their migration feature are XenMotion, vMotion, and KVM migration, respectively. According to our study, we find that there are many factors that affect the performance of the live virtual machine migration such as a long downtime, the large amount of data that is sent in an iteration manner so with a higher dirtying rate the total migration time extends. In our comparison, we show VMWare has the least downtime. We also identify and discuss the various research challenges in detail to stimulate the researchers to work in this direction.

Index Terms: Virtualization, Live virtual machine migration, Pre-copy, Post-copy, Xen, KVM, VMware.

1. Introduction

IT resources have grown more stalwart, more affordable, and can be reached from any place via the internet. Cloud technology is the computing evolution for providing services across the internet. This technology cannot be considered as a fresh concept; after introducing the Amazon EC2 [1], the cloud became a buzz in 2006. It provides applications and everything as a service via the internet using a third individual. The cloud environment has virtualization as its fundamental technology that has been adopted by IBM for underutilized mainframe systems in 1960. Nowadays, this has become the essential technology under numerous modern technologies. It allows the execution of several operating systems on the same host machine, and they can share their resources. The physical host operating system is known as the host, and the guest operating system runs on virtual machines. Virtual machines perform the same task as the physical machines perform. Every system uses a different mechanism to assign physical resources to virtual machines. Virtualization technology provides various services, such as maximum resource utilization, isolation, reliability, better performance, and fault tolerance.

Live migration of virtual machines is a vital virtualization service that migrates on running virtual machines between distinct hosts without detaching the users. Clark et al. [2] first introduced the live virtual machine migration concept. Migration algorithm works in rounds, first, the hypervisor marks dirty to all the memory pages, then its dirty memory pages move iteratively over the network until the quantity of memory pages outlasting or a maximum number of iterations is gotten. After that, hypervisor marks moved pages to clean, because the virtual machine works during live VM migration.

Pre-copy algorithm employs an iterative push stage followed by a stop and copy states. Dirty pages, which are the modified memory pages are generated at the source because of iteration. Memory pages are migrated numerous times as the dirty pages, which are also required to transfer to the destination. Due to it, migration time can be prolonged. It is the first phase, called the pre-copy phase. Memory content is iteratively duplicated page-wise from origin system to

Copyright © 2022 MECS

destination, and VM continues to run. Total pages are migrated in the initial round while VM is still running on the origin host. Few pages may be dirtied and must be retransferred in the next round. The second phase, the termination phase, uses one of the subsequent thresholds; The number of executed repetitions surpasses a predefined threshold, and the total quantity of memory that has already been forwarded surpasses a predefined threshold. The number of pages dirtied in the earlier repetitions falls below the predefined threshold elements, incorporating the applicable criteria that follow. Closing stage is stopped during VM is discontinued and copied the leftover dirty pages and CPU states. When the migration process is finished, the hypervisor as the end-host resumes the virtual machine.

This research performs a comparative performance-based analysis of the Xen [3], KVM [4], and VMware [5] hypervisors, which provide the migration services such as XenMotion, KVM migration, and vMotion, respectively. The essential data transmission techniques can affect live virtual machine migration performance quality. vMotion produces fewer transferred data than XenMotion that offers significant advantages for live migration over local area networks. vMotion acts worse than XenMotion in a specific network with average delay and packet damage, specifying that XenMotion is more appropriate for live migration over a wide area network than VMotion. Accordingly, data center admin practices vMotion for migration in intra data center and XenMotion in inter data center. The quality of vMotion and XenMotion reduce network delay and corruption of packets. The current live migration approach works well in LAN, but this still requires optimization of live migration in wide area network (WAN). The main objectives of this work are shown below:

- To provide a comparative study of the different mechanisms of the live virtual machine migration.
- To critically analyze migration techniques based on the performance metrics and various migration services provided by many hypervisors.
- To provide a brief discussion on the open research challenges that are related to live virtual machine migration in the cloud computing field.

The paper is organized into six sections. Section 2 discusses the different approaches of the live virtual machine migrations and compares them based on the performance metrics. Section 3 presents performance analysis and various performance metrics. The migration functionality provided by various hypervisors is discussed in Section 4. Section 5 discusses the research challenges and future directions, and finally, we conclude our paper in Section 6.

2. Live Virtual Machine Migration

Live virtual machine migration is the method of transferring the virtual machine from one physical server to another without troubling others. The process of copying the virtual machine between physical servers is called migration. The migration approaches can be categorized as static or dynamic. The static mechanism has two methods of migration, whereas dynamic migration has many methods. Static and dynamic methods can be differentiated by their methods of data migration. A virtual machine is fully impracticable when migration is performed using static approaches, whereas dynamic approaches have the least downtime by keeping execution while the migration takes place. The static approaches contain the static method and the cold migration method. In the static migration technique, the VM is fully stopped, and in cold migration, the VM is suspended or frozen in its present state of execution. In live migration or dynamic migration, transfer memory contents between the source host and destination hosts; during this, the VM keeps execution. The motivations behind the live virtual migration are discussed below:

- **Load-balancing:** For load balancing, virtual machines are relocated from overheated or overloaded physical hosts. There is a seldom requirement to shift the virtual machines because of the massive workload on a physical server, virtual machines are moved to reduce the load. Load balancing is done for maintaining services after failure of elements that are observed all the time, then load on system allocated to different systems and from this point onwards sends no traffic to that system.
- **Server consolidation for energy saving:** Server consolidation is the technique of collecting the virtual machines to one host. It is helpful for energy saving of the data centers. In server consolidation, physical servers can be shut down for subsistence, and their workload can be migrated to other physical servers. Virtual machines can be consolidated for energy saving. Few of the underutilized physical hosts shut down their virtual machines, and consolidated physical servers confirm energy efficiency for the green cloud.
- **Online maintenance:** The physical hosts are associated with the customers to enhance the availability and trustworthiness of the system, and maintenance of the host is also essential. To free up the host for maintenance, their load can be transferred to another host. Hence all virtual machines of that physical host can be migrated to another host for sustaining, and services are accessible to clients without interference.
- **The sharing of resources:** The sharing of restricted physical resources like memory, CPU, and cache memory may cause application performance degeneration. This issue can be answered by relocating virtual machines from overburdened hosts to under burdened systems. However, the resources sharing results in diminished operational value due to power-off the useless or idle host.

Copyright © 2022 MECS

2.1 The Standard Live Pre-Copy Virtual Machine Migration Approach

The live migration is a vital service provided by the virtualization software that migrates running VM between separated physical systems without disconnecting with clients. For live migration, the virtual machine state, like the content of the memory and file system, must be transported. The file system can be evaded by employing shared storage. The standard live VM migration performs; first, all the memory pages are listed as dirty, and numbers of repetitions are completed. In the initial repetition, all the memory pages are moved and, in every repetition, the memory pages that get dirtied during the earlier repetition must be re-migrated. If the dirty pages in the origin are under a specific threshold, the virtual machine is halted at the source host. At this period, the state of the processor and the rest of the dirty pages are moved. Lastly, VM is restarted at the target host. Most of the hypervisors like KVM, Xen use the pre-copy mechanism by default, which transfers memory contents and CPU's state.

Sangeeta and Meenu [25] examine the mostly employed precopy approach for VM live migration and propose the two-fold optimization of precopy approach. In the first stage, the amount of data is decreased, which is to be sent in the first round of precopy approach. The second stage limits transferring of identical data iteratively in each successive round by determining repeatedly updated pages and holding them till the final round and copy iteration. As a result, total migration time is decreased.

Pre-copy Standard Algorithm

Step 1: Set_dirty_tracking ()
  For all $(block)_i$ do
    For all $(page)_j$ in $(block)_i$ do
        Set migration dirty flag for $(page)_j$
    End for
  End for
Step 2: Copy all memory pages to destination
Step 3: While (transfer_rate < threshold) do
 Copy memory pages dirtied during the previous iteration
    End while
Step 4: Pause VM
Step 5: Copy the rest of the memory pages and CPU's state
Step 6: Resume VM at the destination

2.2 Post-copy technique

Another mechanism of the live virtual machine migration is post-copy employed by Hines et al. [6]. The post-copy approach contains four essential techniques dynamic self-ballooning, demand paging, pre-paging, and active pushing. The authors have implemented and measured post-copy on Linux-based systems and Xen hypervisor. The evaluations explain that post-copy considerably decreases the total migration time and the transferred pages than pre-copy. The bubbling algorithm for pre-paging can substantially diminish the number of network errors acquired during post-copy migration.

A comparison study of post-copy fronting the pre-copy technique is presented in which Xen hypervisor is employed [7]. This explains advancements in various migration measurements like the total migrated amount of data, complete migration time, and network overhead, which are applying various VM workloads. The authors applied a post-copy approach with adaptive pre-paging to reject all copied page transfers. They remove the transmission of available memory pages in post-copy and pre-copy migration approaches by dynamic self-ballooning. This method regularly issues available pages in virtual machines back to the VMM and accelerates migration with a slight decline in performance.

2.3 Post-copy vs. pre-copy mechanism

Post-copy earliest pauses the target VM at the origin, copies the least CPU state to the destination host, continues the virtual machine, and starts retrieving memory pages across the network from the source. In the pre-copy migration method, first, relocating the memory data to the destination host. After finishing the memory contents transfer, processor states are shifted to the target. Using a post-copy method, the total migration time and the number of pages moved can be diminished compared to pre-copy. While post-copy, a bubbling algorithm is used for pre-paging, diminishing the number of network faults acquired. However, this approach has higher downtime, which prompts service unavailability to the users.

Copyright © 2022 MECS

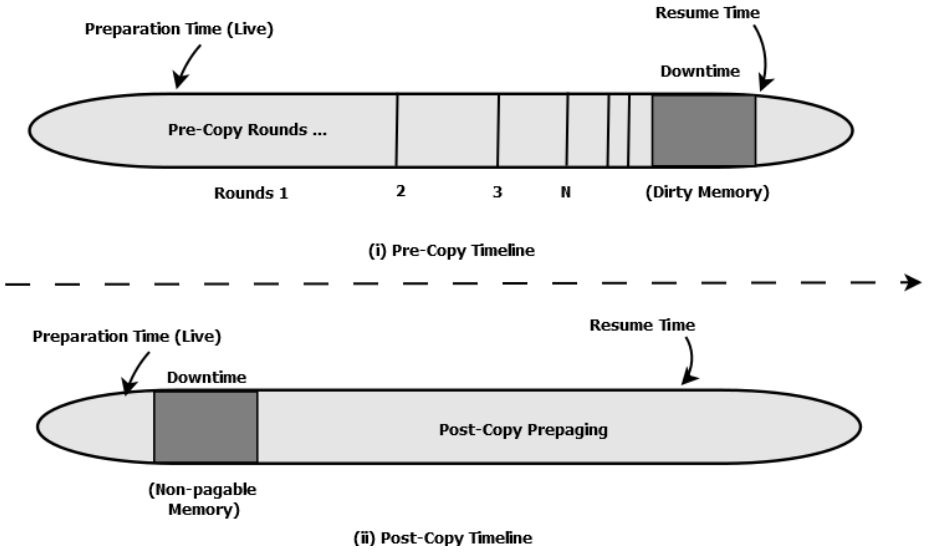


Fig. 1. Pre-copy and post-copy timeline

In the pre-copy approach, total migration time and downtime rely on memory dirtying speed. On the other hand, in post copy approach transfers memory pages only once. Table 1 exhibits a comparison between pre-copy and post-copy methods based on performance metrics. One more difficulty is if any failure befalls during the live VM migration, it cannot be repaired. The total migration time can be estimated as in pre-copy approach and in post-copy approach is as follows:

In pre-copy approach:
Total Migration Time = (size of memory/speed of link) + overhead + non-deterministic dependency on pre-copy rounds (1)

In post copy mechanism:
Total Migration Time = (size of memory/speed of link) + overhead (2)

The total migration time depends on the amount of the data to be transferred, if the amount is larger than it takes a long time as data is transferred in rounds as memory pages get dirtied then again and again transferred and it may cause a long time. The total migration time also inversely depends on link speed, and also other overheads.

3. Performance of Live Virtual Machine Migration

The performance of live migration and non-live are estimated by Kuno et al. [8] and they have shown the performance of processes on a moving VM sternly deterioration. The vital reasons for the deterioration are memory writing and host operating system communication. Garima et al. [24] show an experimental study of live VM migration techniques in KVM in which they analyze the performance and concerning performance metrics total migration time and downtime. They show that downtime during live migration is very smaller than the offline migration, whereas the total migration time is more than the offline migration. We also examine the degradation of the I/O performance reasons. These outcomes show that the leading causes for performance degradation are memory writing and host operating system communication. The performance comparison of vMotion and XenMotion is presented by Feng et al. [9]. vMotion acts better in producing total data of live migration during the migration of virtual machines than XenMotion. vMotion and XenMotion both show the decline in performance in the network with delay and damage of packets. vMotion acts worse than XenMotion in specific networks with average delay and damage of packets. The current live migration methods function adequately in live migration in a small network.

3.1 Performance Metrics

Various researchers discussed several problems in live migration and evaluated the performance based on many performance measures. Voorsluys et al. [10] measure the effect on application performance that is executed on the virtual machine due to the live virtual machine migration using Xen hypervisor. The outcomes demonstrate that the overhead of the live migration is adequate but cannot be ignored, particularly in systems where availability and

responsiveness are administered by Service Level Agreement. To estimate the performance of the virtual machine migration, the parameters [8] are discussed below:

- **Preparation time**: This is the moment live migration has begun and moved the virtual machine's state to the end host. The virtual machine retains execution, and its memory pages get dirtied.
- **Downtime**: This is the duration in which the migrated virtual machines are not live. It involves the transmission of the state of the CPU.
- **Resume time:** During this time, the execution of the virtual machine is continued at the end host, and movement is ended in which all dependencies on the origin host are removed.
- **Pages transferred:** The number of transferred pages, redundant pages across all the duration.
- **Total migration time:** It is the total time of all the durations from starting to stopping. Total migration time is essential as resource deallocation is affected by it.
- **Application degradation:** This is the range to which live migration reduces the execution of the applications in the virtual machine.

3.2 Performance Degradation Factors

While VM is live, its memory gets dirtied at a particular rate, called dirtying rate. If this dirtying rate is larger than network throughput, it increases the number of dirty pages to retransfer. Then VM has to be precipitately discontinued without completing the transferring dirty pages phase. Downtime can be long if there are many numbers of un transferred pages. When a VM machine is migrated over a slow network, then there is a large amount of data to transfer and takes long total migration time. Mostly hypervisor uses a dirty page bitmap, which sets a bit for per dirty page. Through live migration, a bitmap is examined from top to bottom, if the page is listed as dirty, it is sent to the target. In a pre-copy algorithm, memory is transferred in iteration to the destination. In each iteration, pages are dirtied, and they must be resent also. If the dirtying rate is very high, then there is a lot of data for transfer, which causes a long total migration time. Hence, this is a big challenge when pages are updated frequently and are being transferred multiple times during the migration process.

Hypervisors like KVM, Xen use the pre-copy method by default, which transfers memory contents and CPU's state but do not support the transfer of disk contents. Hence if hypervisors like Xen, KVM are being used without changing it, then shared storage such as network-attached storage (NAS) has to be used. Transmission of disk blocks is extremely higher than moving memory and processor states due to a large amount of data. It causes extended total migration time and transferred data, so shared storage is a better option for data centers or cluster environments. VM moves with memory contents in a shared storage environment, and CPU's state and its disk are not transferred from the source node. But if there is no shared disk available and VM is migrated, then disk content must be moved. If the hypervisor is being used without any alterations and the virtual machine is being migrated, then migration could not be completed because of a defect in the migration method. Hence, there must be an algorithm to move VM's disk.

Table 1. Live Virtual Machine Migration Techniques Comparison

Performance Metrics	Pre-copy	Post-copy
Preparation Time	It contains the entire repetitive memory copying phase	Avoidable
Downtime	It involves moving any leftover dirty pages	It involves moving another least execution state
Resume Time	Restart the VM at destination and eliminate the source copy	Most of the post-copy approaches work in this stage
Pages Transferred	It migrates whole pages during preparation time	It transfers most during resume time
Total Migration Time	Higher	Lower
Application Degradation	It must trail dirtied pages by catching write accesses to all pages, which considerably decelerates write-intense workloads	It needs the servicing of significant network errors produced at the destination, which also decelerates VM workloads

4. Live Virtual Machine Migration Services

The various hypervisors provide the live virtual machine facility. There are two main live migration services, XenMotion and vMotion, provided by Citrix [11] and VMware [12], respectively. Both migration services use the pre-copy method for migrating the data, but protocols are different so, they have different performances in the same network condition. The live virtual machine migration services in VMware, Xen, and KVM are discussed below, and a comparison chart is shown in Table 2:

Copyright © 2022 MECS

Table 2. Live VM Migration with different Hypervisor

Hypervisor	Migration Feature	Technique	System Requirement	Process Configuration	Storage	Ethernet Connectivity	Downtime (Approx.)
Xen	XenMotion	Pre-copy	Two physical systems in a resource pool	Same for both system	Shared	Yes	1 or more seconds
VMware	VMotion	Pre-copy	hosts on ESXi 5.1 or later	CPU compatibility	Shared	Yes	zero
KVM	KVM Migration	Pre-copy	QEMU-KVM	Migration possible AMD host to an intel host and back	Shared	Yes	very less

4.1 VMotion in VMware Workstation

VMware is a vendor that offers software for virtualization. The VMware company provides various VMware products such as workstation, VMware Player, VMware Server, and virtual desktop infrastructure. The VMware was stated in 1998 by Diane Greene. VMware established a variety of hypervisors. The Type 1 hypervisor of VMware can execute directly on the hardware. The hypervisor VMware workstation can be executed on x64 systems and on which users can execute various machines with their operating systems such as Linux or Windows. VMware workstation keeps the state of a virtual machine at a particular time. The saved state is called snapshots that can be restored later [19]. VMware workstation executes many virtual machines as one unit that can be powered on, switched off, paused, or resumed as a single unit that can be useful in testing different environments. It offers a secure and efficient platform and supports many operating system such as Windows Server 2008 R2, Windows 7, etc.

vMotion is the service provided by the VMware hypervisor vendor. With the help of it, virtual machines can be migrated from one physical system to another system without stopping. This permits administrators to act on physical server maintenance without disturbing of guest virtual machines that making disaster recovery quicker and in a more effective way. Moreover, vMotion enables load balancing across physical servers, which shows better performance of the virtual machine and enhances resource use. vMotion enables the user to migrate live virtual machines over ESXi hosts without acquiring any downtime. For this purpose, the virtual machine files are required to remain on shared storage. vSphere vMotion allows zero-downtime, live migration of workloads from one physical host to another; hence, users can pursue to access the host they require to stay productive. The vMotion feature of VMware vSphere has zero downtime during the live migration of workloads from one host to another. This ability is feasible across vSwitches, Clouds, and Clusters. While the workload is moving, the application is still working. Users can continue to access the host as they require [20].

Some key features are discussed below:

- **Reliability**: vMotion continues to set the standard for the most reliable live virtual machine migration proficiencies.
- **Performance**: It performs live virtual machine migrations with minimal downtime that is unnoticeable to the customers. The best use of resources such as network and CPU ensures that the live virtual machine migrations happen fast and proficiently.
- **Movements**: vMotion moves across its several hardware generations. VMs are moved from older hosts to newer hosts without downtime or disturbance.
- **Management**: VM migration can be performed easily with the help of a migration wizard, fast recognizing the best target host for a VM that uses real-time information that the migration wizard provides. It also provides several parallel migrations.

4.2 XenMotion

To create the virtualized environment, some virtual platforms, called, the hypervisor is employed such as KVM, Xen, VirtualBox [13] and VMware. Citrix vendor provides Xen hypervisor that is open source. Xen is placed between the physical hardware and the OS that permits several virtual machines to execute concurrently on the same physical host. Each virtual machine works individually, and they have their own assigned region of virtual disks and RAM. Xen hypervisor makes it possible for several guest OS to execute on the same host by adding a virtualization layer that mediates access to the real physical hardware. Xen supports live virtual machine migration that migrates VMs between physical systems over a LAN without disconnecting the users [21]. Throughout this method, the LAN duplicates the memory contents of the VM to the target in the round without breaking its execution and presenting an illusion of

Copyright © 2022 MECS

seamless migration. Citrix Xen has a powerful service XenMotion, which grants capability to the admin to migrate a VM from one physical server to another. VMs migration can be performed from host to host without service disruption for zero downtime. Admins can migrate running application workloads to take benefit of existing computing power.

XenMotion is a service that provides live VM migration. With XenMotion, the running virtual machine can be relocated from one physical system to another, nearly with a small interruption or downtime. XenMotion carries live migration of VMs between hosts with the same type and vendor of the processor. XenMotion is obtainable in all XenServer versions and enables the users to migrate a live virtual machine from one system to another when the virtual machine disks are positioned on storage shared by both systems. This enables pool maintenance characteristics like high availability and Rolling Pool upgrade to migrate virtual machines automatically [22]. These characteristics permit workload levelling, infrastructure flexibility, and server software upgrade without any downtime in virtual machines. XenServer helps for storage XenMotion in version 6.1. Storage XenMotion enables virtual machines to be migrated from one system to another, where the virtual machines are not placed on storage shared between two physical systems. It presents the choice to live migration disks of the virtual machine with the virtual machine itself. It is likely to move a virtual machine from one resource pool to another, move virtual machine disks on local storage, or migrate disks of the virtual machine from one storage container to another.

4.3 KVM

There are many kinds of virtualization, and these are distinct mainly by computing architecture layer. In full virtualization, OS and the applications are executed on the above logical hardware. Kernel-based virtual machine (KVM) is full virtualization that can operate various VMs running unaltered Windows or Linux images. In a live migration, the guest VM stays execution on the source host machine. At the same time, its memory pages are moved to the target host machine. Throughout the migration, KVM observes the source for any modification in pages it has already moved and starts to move these modifications when all the original pages have been moved. KVM also evaluates the moving rate during migration, therefore when the spare volume of data to move will get a specific configurable duration by default 10 milliseconds, KVM pause the source guest VM moves the left data and resumes the same guest VM on the target physical system [23].

5. Research Challenges and Future Directions

For offering continuous services there is the requirement of seamless VM migration. Live VM migration migrates the VMs without cutting off with the customer. For service continuation, live VM migration performance must be extremely high. Modern approaches encounter various difficulties while transferring memory contents, network issues, and data-intensive applications, bandwidth consumption and cloud resources, overloaded virtual machines. Some issues of live VM migration are discussed below:

5.1 VM Migration Across Wide Area Network (WAN)

The current virtual machine migration approaches are not able to deal efficiently with virtual machine migration across a WAN, where the source host and the target host systems are a component of distinct networks [14]. Live VM migration over WAN is a critical issue that may happen due to the many problems that are discussed below:

- **Transferring the connection of storage and network**: - TCP connection persists virtual machine migration and its application and no interruption in network connections if both source and target host belong to the same subnet whereas connection failures in migration occur across subnets.
- **Transferring the storage contents**: - Transfer across a wide area network takes a long because of bulk-sized logical disk files. Therefore, decreasing the volume of data moved across the WAN is a critical task.
- **Persistent state residual at the source host:** - The transferred virtual machine accesses the initial centralized warehouse repository (Storage Area Network) across the wide-area network. Though, because of significant bandwidth utilization and network latencies, it shows poor I/O performance.

5.2 The Speed of Link and Page Dirtying Rate

The link speed is the greatest influential measure for performance. Link potential is inversely proportional to the total migration time and downtime. The rate of getting dirtied pages is also the main factor influencing migration performance. It is the rate at which pages of the memory in the virtual machine are changed that directly influence the number of pages that are migrated in each pre-copy round [15]. Faster migration needs more capable speed links; therefore, it has smaller total migration.

If the pages are getting dirty then rate is more than it grows data sent per repetition, it shows long total migration time and long downtime. The page dirtying rate and the performance of VM migration correlation are not straight. If the dirtied rate of the pages is under link potential, its outcomes have less downtime and total migration time, because of moving all changed pages promptly. Otherwise, the performance of VM migration considerably declines.

Copyright © 2022 MECS

5.3 Virtual Machine Migration with the Heavy Workload

There are various issues in the modern migration approach; the critical issue is when they are employed on heavy workloads like ERP and SAP. They occupy a huge quantity of memory and cannot be migrated as smoothly as lower ones and makes service break shows interferences with heavy applications: interruption of services, disruption of services, hard to maintain consistency and transparency, and unpredictability and rigidity [16].

5.4 Availability of Resources

Resource availability is very significant when a virtual machine is moved. Live virtual machine migration utilizes I/O bandwidths, CPU cycles on the origin and target systems. If there is a requirement of any CPU execution and it is not possible, then migration time would be higher. Therefore, if an important resource is not available, then VM migration could not be performed. Resource availability influences migration performance, such as the influence on the total migration time. Resource availability plays an essential role in live virtual machine migration performance; it can help to make better decisions like moving virtual machines and allocating required resources [17].

5.5 Address Warping

Address-warping issue is one of the problems across wide-area migration. The address of the virtual machine warps from the origin host to the target host makes more difficult of the WAN status and the LANs that linked to the WAN. Hence it is challenging to transfer real-time workload on a virtual machine like conferencing or online games. Due to it, long downtime can occur, so long downtime and complications must be evaded [18].

6. Conclusion

Cloud computing has virtualization as its key technology and there are many reasons for using it that can be employed for one physical system in which various platforms can be used for several purposes. Virtualization technology is the key concept of various modern environments like cloud computing. Live virtual machine migration is one of the several benefits provided by virtualization. This research discusses the various migration services which are provided by the many hypervisors. We show the performance analysis of these migration services. There are various performance parameters that are affected by many factors like link speed, and the size of applications that are run on virtual machines, VM migration over WAN. This research discusses various research challenges, which can be useful to provide a future direction to the other researchers.

References

[1] 'Amazon Elastic Compute Cloud (Amazon EC2)', https://aws.amazon.com/ec2/. 2021.
[2] Clark C., Fraser K., Hand S., Hansen J. G., Jul E., Limpach C., Pratt I., and Warfield A., 'Live migration of virtual machines', In Proceedings of the 2nd Conference on Symposium on Networked Systems Design & Implementation, 2005, Vol. 2, pp. 273-286.
[3] 'Xen Hypervisor', https://www.xen.org/products/xenhyp.html. 2021.
[4] 'Kernel-based Virtual Machine', https://linux-kvm.org. 2021.
[5] 'vSphere ESX and ESXi Info Center', https://vmware.com/products/vsphere/esxi-and-esx. 2021.
[6] Shribman, A., and Hudzia, B., 'Pre-Copy and Post-Copy VM Live Migration for Memory Intensive Applications', In: et al. Euro-Par 2012: Parallel Processing Workshops. Euro-Par 2012. Lecture Notes in Computer Science, Vol 7640. Springer, Berlin, Heidelberg.
[7] Sarddar, D. and Nandi, E., 'Efficient Virtual Machine Migration with Reduced Migration Time and Downtime', International Journal of Urban Design for Ubiquitous Computing. Vol. 3, 2015, pp. 1-8.
[8] Kuno, Y., Nii K., and Yamaguchi, S. 'A Study on Performance of Processes in Migrating Virtual Machines', Tenth International Symposium on Autonomous Decentralized Systems, 2011, pp. 567-572.
[9] Feng, X., Tang, J., Luo, X., and JinX Y., 'A performance study of live VM migration technologies: vMotion vs XenMotion', In Asia Communications and Photonics Conference and Exhibition, Optical Society of America, 2011, pp. 83101.
[10] Voorsluys, W., Broberg, J., Venugopal, S., and Buyya, R. Cost of Virtual Machine Live Migration in Clouds: A Performance Evaluation. In: Jaatun, M.G., Zhao, G., Rong, C. (eds) Cloud Computing, CloudCom 2009. Lecture Notes in Computer Science, Vol 5931. Springer, Berlin, Heidelberg.
[11] 'Citrix', http://www.citrix.com/products/. 2021.
[12] 'VMware', http://www.vmware.com. 2021.
[13] 'VirtualBox', https://www.virtualbox.org/. 2021.
[14] Bose, S. K., Brock, S., Skeoch R., and Rao, S., 'CloudSpider: Combining Replication with Scheduling for Optimizing Live Migration of Virtual Machines across Wide Area Networks', 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2011, pp. 13-22.
[15] Akoush, S., Sohan, R., Rice, A., Moore, A. W., and Hopper, A., 'Predicting the Performance of Virtual Machine Migration', IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, 2010, pp. 37-46.

Copyright © 2022 MECS

[16] Hacking, S., and Hudzia, B., 'Improving the live migration process of large enterprise applications', In Proceedings of the 3rd International Workshop on Virtualization Technologies in Distributed Computing, 2010, pp. 51-58.
[17] Wu, Y. 'Performance modelling of virtual machine live migration', In IEEE 4th International Conference on Cloud Computing, IEEE, 2011, pp. 492-499.
[18] Kanada, Y., and Tarui, T., 'A "network-paging" based method for wide-area live-migration of VMs', The International Conference on Information Networking (ICOIN2011), 2011, pp. 268-272.
[19] Gupta, V. K., Rana, P. S., 'Toxicity prediction of small drug molecules of androgen receptor using multilevel ensemble model, Journal of Bioinformatics and Computational Biology', 2019, Vol. 17, pp. 1–26.
[20] Jaiswal, N., Gupta, V. K., & Mishra, A., Survey paper on various techniques of recognition and tracking. In 2015, International Conference on Advances in Computer Engineering and Applications, IEEE, 2015, pp. 921-925.
[21] Yadav, P., Varshney, R., and Gupta, V. K., Diagnosis of breast cancer using decision tree models and SVM. International Research Journal of Engineering and Technology (IRJET), 2018, pp. 2395-0056.
[22] Gupta, V. K., Surendra Kumar Shukla, Anupriya, & Ramesh Singh Rawat, Crime Tracking System and People's Safety in India using Machine Learning Approaches. International Journal of Modern Research, Vol. 2, 2022, pp. 1–7.
[23] Shukla, S. K., Gupta, V. K., Joshi, K., Gupta, A., and Singh, M. K. (2022). Self-aware Execution Environment Model (SAE2) For the Performance Improvement of Multicore Systems. International Journal of Modern Research, 2022, Vol. 2, pp. 17-27.
[24] Rastogi, G., and Sushil, R., 'Performance analysis of live and offline VM migration using KVM', International Journal of Modern Education and Computer Science, 2016, Vol. 11, pp. 50-57.
[25] Sharma, S., and Chawla, M., 'Two Fold Optimization of Precopy Based Virtual Machine Live Migration', International Journal of Information Technology and Computer Science, 2015, Vol. 10, pp. 8-14.

Authors' Profiles

Divya Kapil is working as an Assistant Professor in Graphic Era Hill University, Dehradun. She has received M.Tech degree in Computer Science and Engineering from Graphic Era Deemed to Be University, Dehradun, India in 2013. She is currently pursuing Ph. D from Graphic Era Deemed to be University, Dehradun, India. Her research interest includes Security in cloud computing and virtualization.

Dr. Vishan Kumar Gupta has received his Ph.D. degree in Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India. He has received his M.Tech degree in Information Technology from ABV-Indian Institute of Information Technology and Management, Gwalior, MP, India and Bachelor of Engineering in CSE from Rajiv Gandhi Technical University, Bhopal, MP, India, Currently, he is working as an Associate Professor in Graphic Era Deemed to be University, Dehradun, Uttarakhand, India. His areas of research are Computational Biology, Data Mining, and Machine learning.

Saurabh Kumar Mishra is working as an Assistant Professor in Graphic Era Deemed to be university, Dehradun. He is M. Tech from Delhi Technological University (formerly DCE). His area of research is cloud computing and machine learning.

How to cite this paper: Divya Kapil, Saurabh Kumar Mishra, Vishan Kumar Gupta, "A Performance Perspective of Live Migration of Virtual Machine in Cloud Data Center with Future Directions", International Journal of Wireless and Microwave Technologies(IJWMT), Vol.12, No.4, pp. 48-56, 2022. DOI:10.5815/ijwmt.2022.04.04

View publication stats

Arabian Journal of Geosciences (2022) 15:1331
https://doi.org/10.1007/s12517-022-10567-8

ORIGINAL PAPER

A picture of Delhi's regional air quality during diminished anthropogenic activities in the COVID-19 era

Vignesh Mohan[1] · **Rajeev Kumar Mishra**[1]

Received: 5 January 2022 / Accepted: 30 June 2022
© Saudi Society for Geosciences 2022

Abstract

COVID-19 pandemic lockdown was imposed in 2020 and 2021 globally and in India, which created a temporary switch off for various anthropogenic activities, which helped in improving Delhi's air quality drastically. The present study discusses air pollution reduction during that lockdown period compared to the previous year. The study tries to establish a comparison between various parameters like particulate matter (PM_{10}, $PM_{2.5}$), gaseous pollutants (SO_2, NO_2, CO, O_3), and meteorological data (WS, WD, RH, AT, SR) during the year 2019 to 2021. The whole monitoring period is divided into four phases (P-I, P-II, P-III, and P-IV) based on stage-wise lockdown implementation. The data are analyzed using the R Studio version 1.3.1093. The PM_{10} concentration falls to 46.6% and 19.2% in 2020 and 2021, compared to 2019, and 69% increase in 2021 from 2020. $PM_{2.5}$ reduced to 35% and 8.1% in 2020 and 2021, compared to 2019, and a 48% increase in 2021 from 2020. The O_3 concentration shows an increment of about 10% in 2020 and a reduction of 5% in 2021, compared to 2019, and 12%, compared to 2020. The SO_2 and CO reductions lie in between 20 and 35% in 2020 and 2021, compared to 2019. The NO_2 shows around 56% and 32% reduction during 2020 and 2021 compared to 2019 and a 62% increment in 2021 compared to 2020. An absolute lockdown was imposed in 2020, whereas minimal activity or certain restrictions were only imposed in 2021. Such changes in concentration show that anthropogenic sources played a significant role in determining regional air quality in urban areas. The outcomes of the study show the changes that occurred during that period due to phase-wise implementation of lockdown and other unlocking conditions, which may be helpful for policymakers in implementing certain restrictions based on emissions in a planned manner in the future.

Keywords Anthropogenic activities · COVID-19 · Lockdown · Particulate matter · Urban air pollution

Abbreviations

AT	Atmospheric temperature
AQI	Air quality index
AY	Actual year
BC	Black carbon
BS-IV	Bharat stage four
BS-VI	Bharat stage six
BY	Base year
$\mu g/m^3$	Microgram per meter cube
CPCB	Central Pollution Control Board
CO	Carbon monoxide
CoI	Cost of illness
COVID-19	Coronavirus disease
DALYs	Disability-adjusted life years
Hrs	Hours
ICMR	Indian Council of Medical Research
IDW	Inverse distance weighting
IQAir	World Air Quality Report
mg/m^3	Milligram per meter cube
m/s	Meter per second
NASA	National Aeronautics and Space Administration
NCR	National Capital Region
NCT	National Capital Territory
NH_3	Ammonia
NO	Nitric oxide
NOx	Nitrogen oxide
NO_2	Nitrogen dioxide
O_3	Ozone
P-I	Phase I
P-II	Phase II

Responsible Editor: Amjad Kallel

✉ Rajeev Kumar Mishra
rajeevkumarmishra@dtu.ac.in

[1] Department of Environmental Engineering, Delhi Technological University, Delhi 110042, India

Published online: 25 July 2022

P-III	Phase III
P-IV	Phase IV
PAHs	Polycyclic aromatic hydrocarbons
PC	Pearson correlation
PC	Percentage change
PCI	Per capita income
Per sq. km	Per square kilometer
$PM_{2.5}$	Particulate matter with a diameter of 2.5 µm
PM_{10}	Particulate matter with a diameter of 10 µm
RH	Relative humidity
SO_2	Sulfur dioxide
SR	Solar radiation
VOCs	Volatile organic compounds
VSL	Value of statistical life
WD	Wind direction
WS	Wind speed
WHO	World Health Organization

Introduction

Many countries face globally tremendous issues to mitigate the global pandemic COVID-19, and still, many countries are facing the same (Sikarwar and Rani 2020; Li et al. 2020). The new infectious (COVID-19) disease was first identified in Wuhan City of China in December 2019 (Gautam et al. 2021b; He et al. 2020; Srivastava et al. 2020). The World Health Organization (WHO) in January 2020 declared that the virus could be transferred to humans as a droplet infection and declared a health predicament throughout the globe due to this infectious disease (Sarfraz et al. 2020; Bera et al. 2021). In India, the first case was reported on 30 January 2020 in Kerala, a southern state of the country, and it declared an emergency on 4 February 2020 since another case was reported on the fourth consecutive day (Gautam et al. 2021a). Until the 2nd of March 2020, the country had only three reported cases, but the number of human contacts has increased tremendously in April 2020 (Sarfraz et al. 2020). To control this pandemic, many countries have implemented lockdown.

Similarly, in India also, lockdowns were implemented (Bherwani et al. 2020; Dhaka et al. 2020). In Delhi, the state government implemented a night curfew from March 22, 2020 (19:00 Hrs to 07:00 Hrs). Shortly after that, on March 24, a nationwide lockdown was enforced. Phase I of the lockdown was announced for 21 days from March 25 to April 14, 2020. Furthermore, Phase II was implemented from April 14 to May 3, 2020. During this period of Phases I and II, all activities such as commercial, industrial, and transportation activities at all levels came to a pause except the basic amenities such as medical and food. The Oxford COVID-19 Government Response Tracker report stated that India had one of the strictest implementations of lockdown. Phase III of the lockdown was announced nationwide from May 4 to 17, 2020, giving few relaxations such as the operation of essential government services with 50% occupancy. The Phase IV nationwide lockdown was announced on May 17 and extended until May 31, 2020, with relaxations such as the operation of shops and other essential services during the daytime. After that, the country started to give several relaxations based on the infection rate in a particular location and then the lockdown scenarios were slowly reduced stepwise such as Unlock 1, 2, and 3 and the conditions were resuming back to normal with certain restrictions such as maintaining social distancing and other hygiene practices at the places where the crowd is expected and the scenario continues until 2021, but unfortunately in 2021, the second wave of the pandemic hit harder than the first wave in India. Many Indian cities, including Delhi, had experienced the worst situation, and the patients were reported in large numbers. Even the hospitals in urban regions were running out of beds for patients, and oxygen for medical use was in high scarcity. Due to this, the government of Delhi announced another set of week-wise lockdowns starting from 19 April to 31 May 2021. During this period, minimal movement of people and vehicles was observed. Compared to the movement scenario in 2020, it was a little bit higher in 2021. Despite several issues reported due to pandemic such as economy and human health, it was vice versa for pollution related issues. In 2020 several research studies showed a reduction in pollution levels in Delhi, which reflected improved air quality (Karuppasamy et al. 2020; Agarwal et al. 2020; Sikarwar and Rani 2020; Mahato et al. 2020; Hamaa et al. 2020; Srivastava et al. 2020; Kumar et al. 2020). All types of transportation services, including rail, road, air, and all industrial units except essential services, were suspended in 2020, leading to a dramatic reduction in air pollution (Kumari and Tosniwal 2020), but in 2021 there were exceptions given for movement.

The World Urbanization Prospect in 2018 predicted that the megacities of Asia and Africa are about to experience 90% population growth by 2050 (United Nations 2019). Air pollution has become a critical threat to the environment and causes serious threats to human health. About 80% of the people in urban areas are exposed to air pollution exceeding the air quality standard set by the World Health Organization (WHO). As per the World Air Quality (IQAir) (WORLD AIR QUALITY REPORT 2019) report, six major Indian cities with poor air quality were recorded globally. Delhi has consistently been ranked as one of the most polluted cities in developing countries (Sikarwar and Rani 2020). The sources of air pollution in Delhi include power generation, industry, traffic, household activities, pavement dust,

and biomass burning (Bherwani et al. 2020; Venter et al. 2020). Due to poor air quality, a community health emergency was declared in 2017 for Delhi by the Indian Council of Medical Research (ICMR) (Sikarwar and Rani 2020). To tackle this situation, several short-term and long-term measures were taken to improve the air quality. The short-term remedies such as the odd–even scheme were introduced in 2016 and 2019 when the air quality of Delhi worsened and the levels were extremely high compared to the standards. It showed a particular percentage reduction of pollutants in Delhi (Khaiwal et al. 2021; Mishra et al. 2019; Kumar et al. 2021). Long-term measures include policy changes such as conversion from BS-IV to BS-VI, scrappage of older vehicles, and other policy changes. Global studies found that reducing motor vehicle usage positively affects air quality (Higham et al. 2021). Several studies in the year 2020 found a reduction in air pollution globally due to the lockdown scenarios, and in Indian scenarios, it was also the same, especially the national capital experienced good air quality due to the reduced anthropogenic activities. In Delhi, the complex mixture of the different sources of pollutants is the biggest problem, and during that lockdown period, all the different sources were switched off, which helped in improving air quality. The previous studies during the lockdown period (2020 only) showed how much reduction happened due to the lockdown. This study aims to bring out the changes that occurred not only during the lockdown period (2020) but also during limited sources activity (2021) compared with the usual sources scenario (2019). The objectives of the study is to identify the percentage changes that occurred in Delhi due to the lockdown and series of restriction phases and also to compare it with the previous year. The study utilizes around forty locations data to represent the study area which will be quite helpful for the representation. The outcomes of the study can be used as representative data for the study area since it covers data from all the different locations possible. The study provides information in different scenarios of Delhi along with different sources. The changes that occurred will help in future studies of the location in acting as a baseline measurement of less anthropogenic sources and the limitation of the study is that it compared the data during that particular period only when there was a lockdown or restriction. A summary of a few recent research works published based on COVID-19 lockdown and its impacts on the air quality in Indian cities are depicted in supplementary Table S1.

Methodology

The time series data of various primary pollutants and meteorological parameters are taken from the time period of 25 March to 31 May 2020 for 2019, 2020, and 2021. The period is classified into four phases, namely, Phases I, II, III, and IV, which were classified based on the stepwise lockdown implemented by the government of India in the year 2020. The Phase I (P-I) was from 25 March 2020 to 14 April 2020, Phase II (P-II) from 15 April to 3 May 2020, Phase III (P-III) from 4 to 17 May 2020, and Phase IV (P-IV) from 18 to 31 May 2020. The comparison was made for 2019 with 2020 (when there was a strict lockdown), 2020 and 2021 (when there was a restriction but not a strict lockdown), and 2019 and 2021 (from normal period to minimal anthropogenic activities). This analysis presented a clear picture of the air quality in these 3 consecutive years due to the variation in quantity and quality of sources during the different eras of the COVID-19 pandemic. This comparative study helped to see the effect of lockdown/movement on Delhi's air quality. The data sets are taken from the Central Pollution Control Board (CPCB) open source. The pollutants (PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, O_3) and meteorological parameters (wind speed, wind direction, temperature, relative humidity, and solar radiation) are taken from the CPCB website (https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/data).

Selection of the study area

The national capital of India, Delhi, has been selected to conduct this study. It is selected for this study since Delhi is among the largest megacities globally, the national capital, and the single largest contributor to India's urban population (about 7.6%). The land cover is about 1485 km^2 area with a population distribution of about 16.8 million. In the past two decades, the population density has increased from 9340 people per sq. km in 2001 to 11,297 people per sq. km in 2011, with an increased rate of 37.6% annually (Mahato et al. 2020). A total of 40 operational monitoring stations in NCR and NCT part of Delhi were taken into consideration for this study (Fig. 1). The different sources of the data providers are given in supplementary Table S2. The various locations selected for extracting the data are provided in supplementary Table S3.

Analyses technique

The data are analyzed using the RStudio version 1.3.1093. Hourly average data were taken from 25 March to 31 May for 3 consecutive years, from 2019 to 2021. The temporal variation of the pollutants during the different phases was taken and a comparison was made between the years 2019 and 2020, 2020 and 2021, and 2019 and 2021. The data set has been divided into four different phases (P-I, P-II, P-III, P-IV) based on the

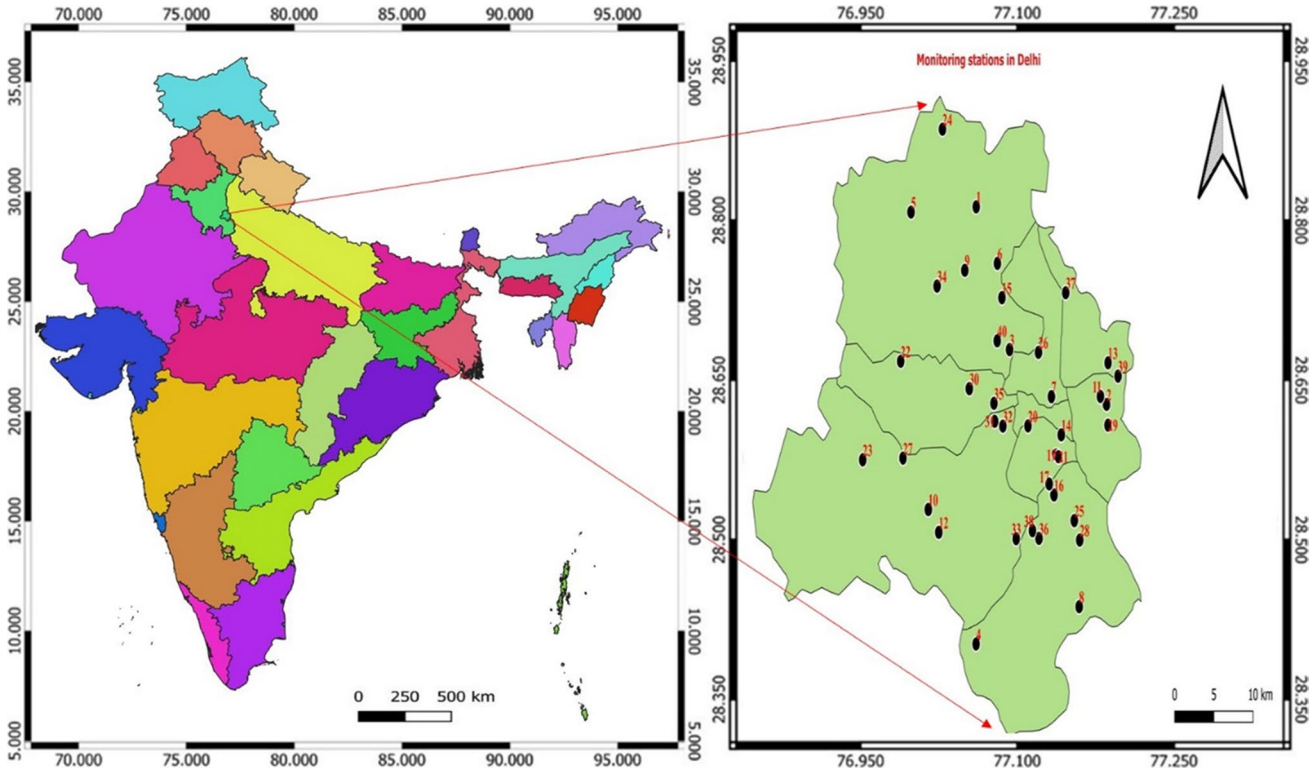


Fig. 1 Map displaying the monitoring stations

lockdown implementation as per 2020 government guidelines and simultaneously compared with 2019 and 2021. The percentage change has been calculated using the following equation:

$$PC = \frac{AY - BY}{BY} * 100 \quad (1)$$

where *PC* is the percentage change that occurred, *AY* refers to the average concentration of a particular pollutant in the actual year taken for comparison, and *BY* refers to the average concentration of a particular pollutant in the base year.

Result and discussion

Bar plot for phase-wise percentage change

The bar plot shows the percentage change that transpired during 2020, taking 2019 as a BY and 2020 as the AY (Fig. 2). It is very clear from the bar plot that the concentration of all the parameters is reduced due to the enactment of strict lockdown. Interestingly, the concentration of O_3 increased in P-II, P-III, and P-IV compared to the

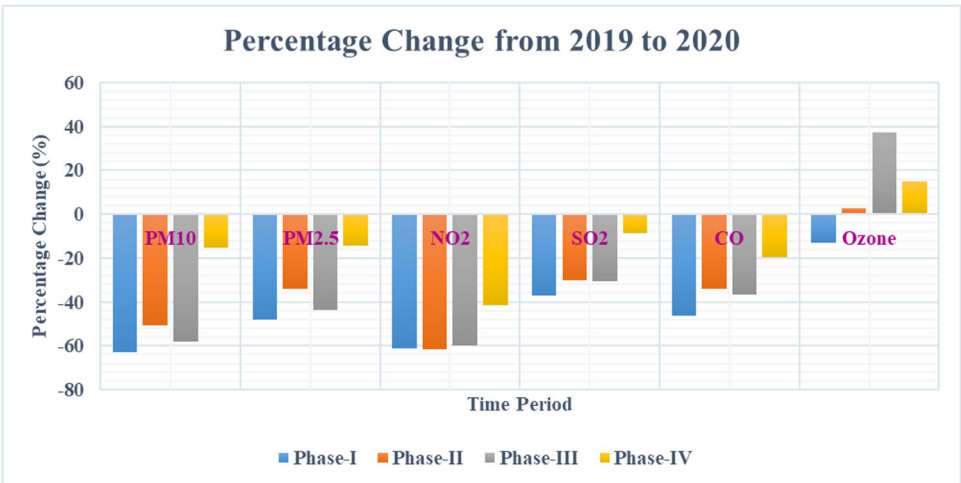


Fig. 2 Pictorial representation of percentage change occurred in 2020 compared to 2019 for PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, and O_3 in Phases I, II, III, and IV

BY to a maximum concentration of 62.3 μg/m^3 with a percentage increment of 37.2% compared to the normal scenario. The O_3 concentration increased during the strict lockdown period, i.e., in 2020. (Naqvi et al. 2021; Gautam et al. 2021b). The particulate matter (PM_{10}) showed a significant percentage reduction (Srivastava et al. 2020) in P-I, P-II, and P-III and less reduction in P-IV, which may be due to the increased anthropogenic activity on account of various relaxation. The minimum concentration of PM_{10} recorded was 84.1 μg/m^3. $PM_{2.5}$ concentration followed the same trend as that of PM_{10} and the minimum recorded concentration was 34.9 μg/m^3 with an average reduction of 35% in all four phases. During the lockdown period, the decrease in $PM_{2.5}$ concentration was also around 41–53% (Kumar et al. 2020). The declines in SO_2 (Kumari and Toshniwal 2020) and CO concentrations were also seen (Sarfraz et al. 2020; Gautam et al. 2021b; Sikarwar and Rani 2020). The percentage reduction of pollutants was seen high during P-I, P-II, and P-III and less in P-IV (Bherwani et al. 2020; Vadrevu et al. 2020; Hamaa et al. 2020; Khaiwal et al. 2021; Mahato et al. 2020). The phase-wise overall percentage changes are depicted in Table 1. Recent researchers have found that 79% of PM_{10} emissions originate from Delhi road dust and resuspensions (Kumar et al. 2020).

The concentrations of PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 were recorded high in 2021 compared to 2020 in all phases except P-IV because this particular phase (P-IV) may have the same amount of source contribution. In 2020, certain relaxation was given from strict lockdown during P-IV, and in 2021, in all phases, certain restrictions were only announced, and it was not a complete lockdown in the city (Hamaa et al. 2020). In Fig. 3, the percentage change was calculated using Eq. 1, where BY is 2020 and AY is 2021. The percentage change in P-I for parameters PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 is 167%, 104.3%, 142.3%, 42.9%, 70%, and 6.8% respectively. Similarly for P-II, it is 109.9%, 106.6%, 92.3%, 28.9%, 26.5%, and − 7.2%. The percentage change for other phases can be observed in Table 1.

Figure 4 depicts exciting results that the concentration of PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 is reduced except for some exceptional phases like P-II in the year 2021 compared

Table 1 Phase-wise percentage change calculation

Phase	BY	AY (% change)	PM_{10} (μg/m^3)	$PM_{2.5}$ (μg/m^3)	NO_2 (μg/m^3)	SO_2 (μg/m^3)	CO (mg/m^3)	Ozone (μg/m^3)
Phase I	2019	2020	− 62.85	− 48.11	− 61.27	− 37.08	− 46.31	− 13.19
	2020	2021	167.09	104.32	142.37	42.91	70.08	6.83
	2019	2021	− 0.78	6.01	− 6.12	− 10.08	− 8.69	− 7.26
Phase II	2019	2020	− 50.60	− 34.27	− 61.71	− 29.90	− 34.24	2.65
	2020	2021	109.91	106.15	92.31	28.91	26.56	− 7.28
	2019	2021	3.69	35.51	− 26.36	− 9.64	− 16.77	− 4.83
Phase III	2019	2020	− 58.11	− 43.61	− 59.83	− 30.43	− 36.55	37.29
	2020	2021	18.43	4.87	27.86	− 12.61	18.88	− 20.58
	2019	2021	− 50.39	− 40.86	− 48.64	− 39.21	− 24.57	9.04
Phase IV	2019	2020	− 15.22	− 14.39	− 41.36	− 8.60	− 19.53	14.75
	2020	2021	− 16.76	− 21.98	− 13.86	− 45.54	− 13.88	− 29.65
	2019	2021	− 29.43	− 33.21	− 49.49	− 50.22	− 30.69	− 19.27

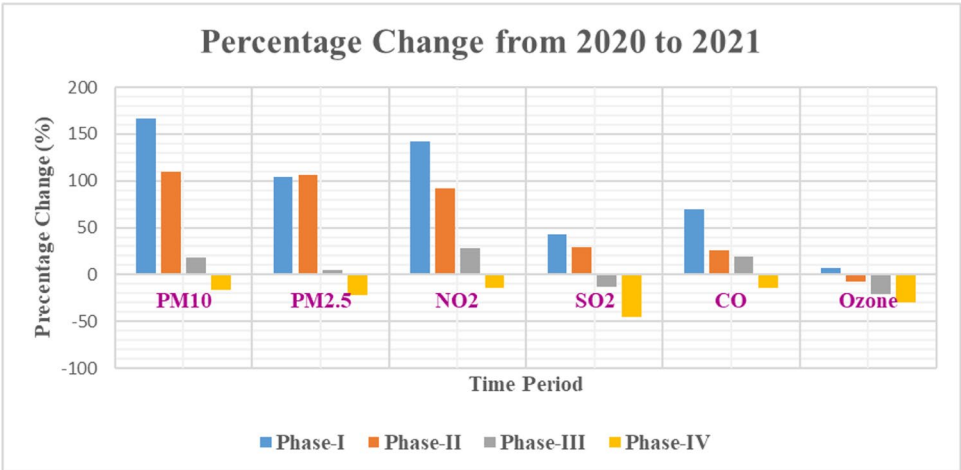


Fig. 3 Pictorial representation of percentage change occurred in 2021 compared to 2020 for PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, and O_3 in Phases I, II, III, and IV

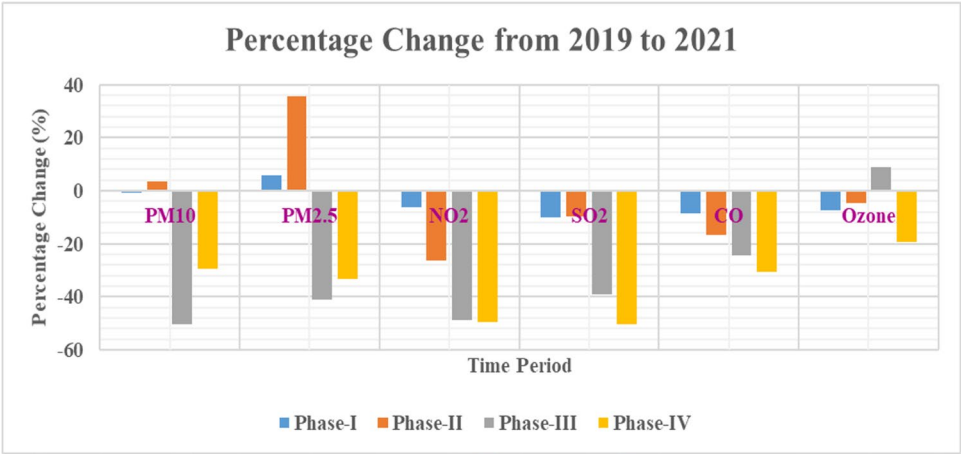


Fig. 4 Pictorial representation of percentage change occurred in 2021 compared to 2019 for PM_{10}, $PM_{2.5}$, SO_2, NO_2, CO, and O_3 in Phases I, II, III, and IV

with the usual scenario of 2019. The year 2019 was a normal year in which all the activities were functional. In 2021, certain restrictions were amended and the sources were also restricted. But the lockdown implemented in 2021 was not an absolute switch off to all the sources. The minimum concentration of PM_{10} and $PM_{2.5}$ during 2021 was 131 μg/m^3 and 45.8 μg/m^3, respectively, whereas the maximum was 235.9 μg/m^3 and 93.9 μg/m^3, accordingly. The maximum concentration reduction for NO_2, SO_2, CO, and O_3 was seen in P-IV in 2021. The gradual phase-wise reduction of NO_2, SO_2, CO, and O_3 can be seen clearly in Fig. 3. Equation 1 has been used for percentage change calculation for BY and AY. The exact figures of the increment or decrement in percentage can be obtained from Table 1. It shows that anthropogenic activities play a significant role in determining the air quality of a particular region (Higham et al. 2021; Venter et al. 2020).

Temporal representation of concentration in different phases

The lockdown impact on the hourly mean concentration of PM_{10} was studied for 3 consecutive years from 2019 to 2021.

The values of PM_{10} tend to decrease in 2020 and 2021 compared to 2019. The average mean concentration in 2019 was 226.4 ± 81.8 μg/m^3 (P-I), 227.5 ± 101.2 μg/m^3 (P-II), 292.2 ± 135.6 29 μg/m^3 (P-III), and 185.6 ± 77.7 μg/m^3 (P-IV), respectively. It reduced to 84.1 ± 38.5 μg/m^3, 112.4 ± 52.3 μg/m^3, 122.4 ± 52.5 μg/m^3, and 157.4 ± 76.1 μg/m^3 in 2020, and also similarly, there was an increment in 2021 when compared to 2020, but the concentrations in

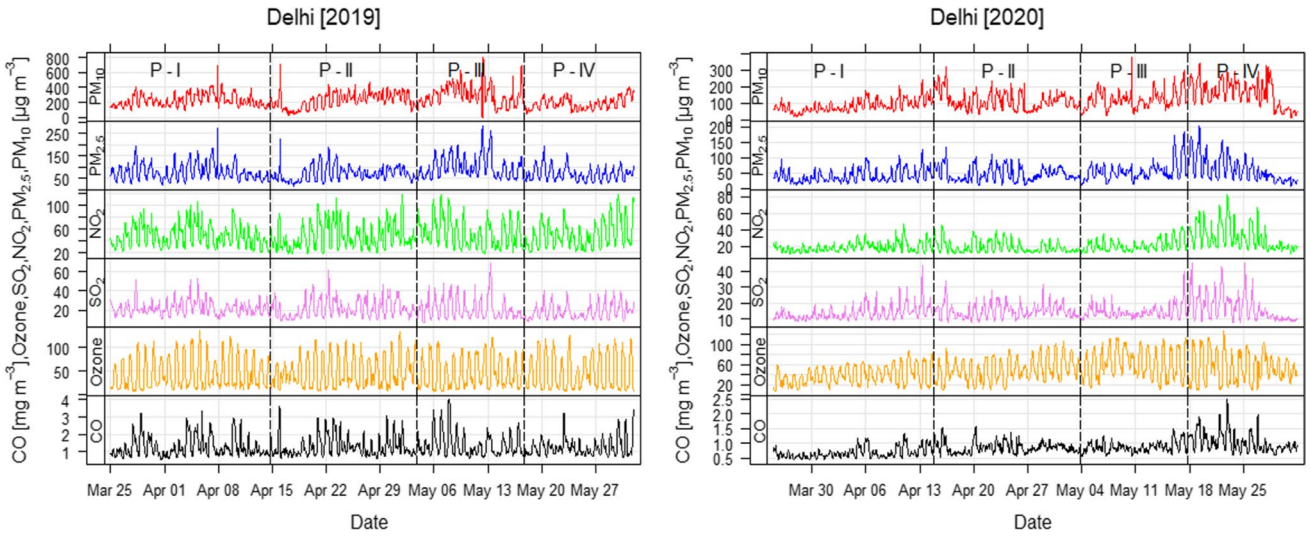


Fig. 5 Hourly average concentration of pollutants PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 in Phases I, II, III, and IV for 2019 and 2020

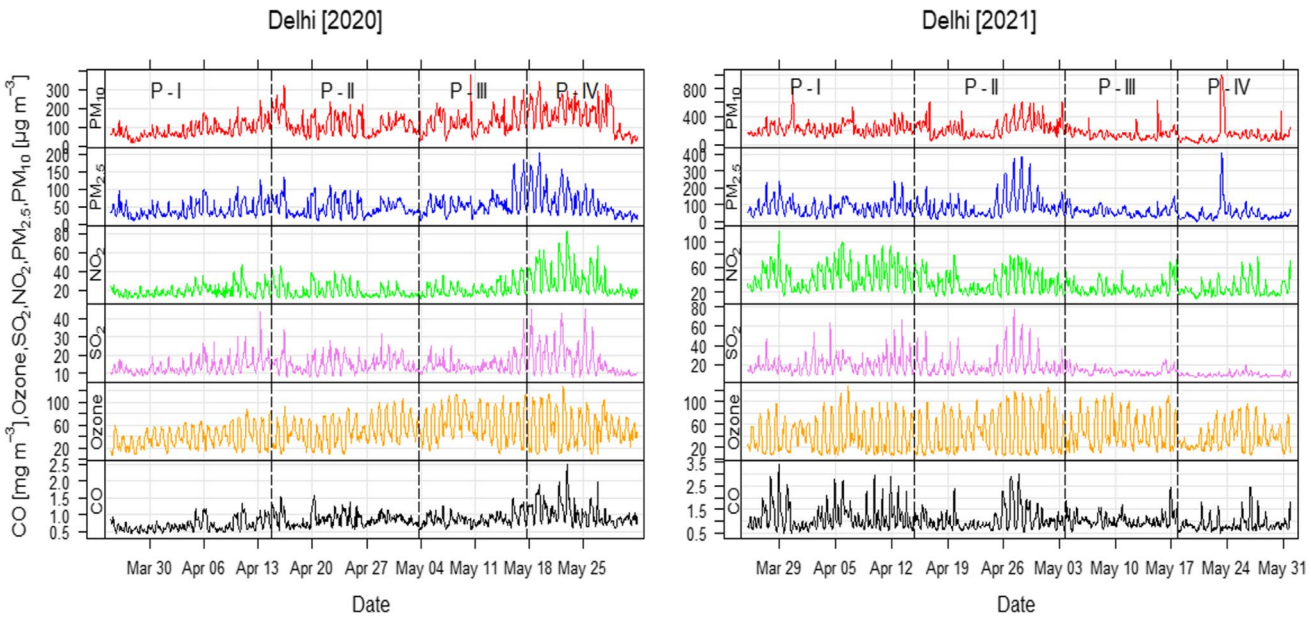


Fig. 6 Hourly average concentrations of pollutants PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 in Phases I, II, III, and IV for 2020 and 2021

both the year 2020 and 2021 are lesser than the concentration recorded in the year 2019, but in the P-III and P-IV of 2021, they show reduction due to gradual implementation of lockdown process during this phase (e.g., the average mean concentration of PM_{10} was 157.4 ± 76.1 µg/m^3 in P-IV of 2020 but in 2021 it was 131 ± 147.2 µg/m^3). Figure 5 depicts the daily average concentration of all the parameters, and it shows that in P-III and P-IV, there is a gradual increase in concentration due to stepwise implementation of the unlocking procedure, and the maximum concentration of pollutants was observed during P-IV of 2020.

In P-IV, the concentration decreases to a minimum level in 2021 (Fig. 6). It indicates that the implementation of lockdown during the time in 2021 influenced the pattern of the curve. In Fig. 7, all the four phases, namely, P-I to P-IV, show a lower concentration of pollutants in 2021 compared to 2019. It shows that the year 2019 has no bound for activities, but in 2021, certain limitations in different

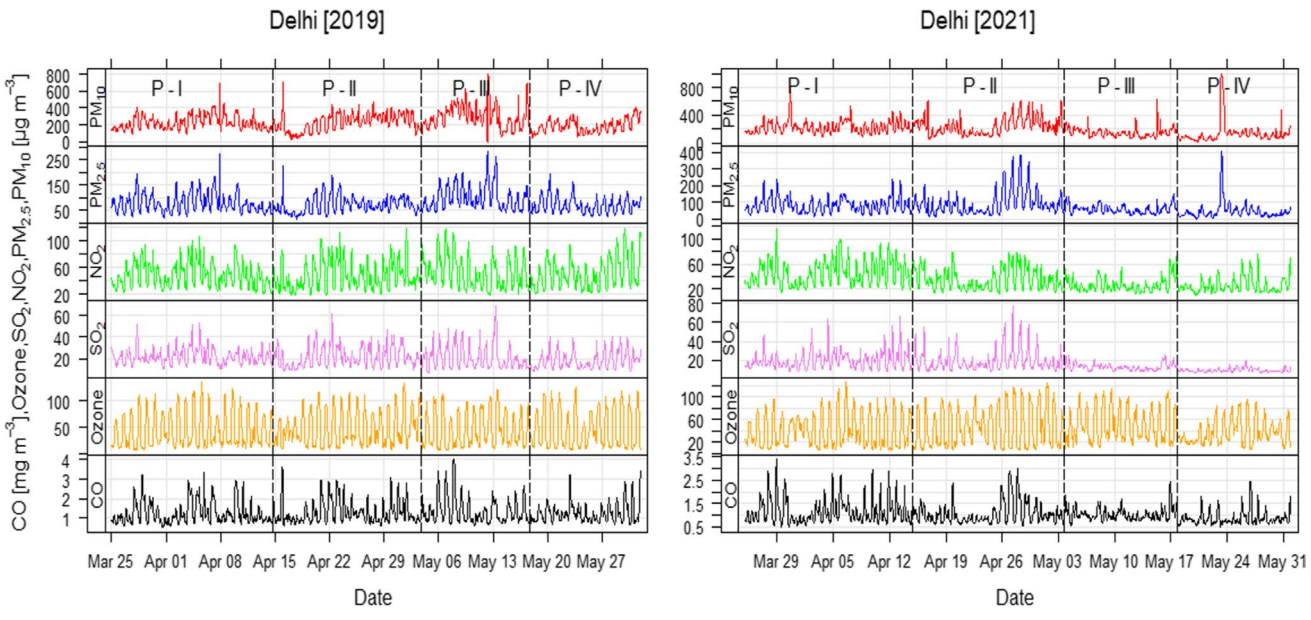


Fig. 7 Hourly average concentrations of pollutants PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 in Phases I, II, III, and IV for 2019 and 2021

Table 2 Summary statistics of hourly average concentration of pollutants

	Year	2019		2020		2021	
Phase	Parameter	Mean ± SD	Median	Mean ± SD	Median	Mean ± SD	Median
Phase I	**PM_{10} ($\mu g/m^3$)**	226.49 ± 81.87	209.67	84.14 ± 38.50	75.07	224.73 ± 87.90	209.9
	$PM_{2.5}$ ($\mu g/m^3$)	76.92 ± 36.64	68.76	39.91 ± 20.55	32.83	81.55 ± 41.35	76.62
	NO_2 ($\mu g/m^3$)	48.38 ± 17.95	45.38	18.74 ± 5.89	17.31	45.42 ± 19.56	43.09
	SO_2 ($\mu g/m^3$)	21.61 ± 6.85	20.65	13.59 ± 4.34	12.371	19.43 ± 8.87	17.6
	CO (mg/m^3)	1.26 ± 0.56	1.0682	0.67 ± 0.18	0.6217	1.15 ± 0.51	0.9842
	O_3 ($\mu g/m^3$)	44.38 ± 32.13	28.571	38.53 ± 18.09	35.753	41.16 ± 29.44	32.358
Phase II	**PM_{10} ($\mu g/m^3$)**	227.59 ± 101.29	232.17	112.42 ± 52.33	102.86	235.98 ± 120.79	211.73
	$PM_{2.5}$ ($\mu g/m^3$)	69.31 ± 32.23	62.33	45.56 ± 23.57	38.509	93.92 ± 71.26	73.58
	NO_2 ($\mu g/m^3$)	47.99 ± 21.70	41.75	18.37 ± 6.87	15.54	35.34 ± 16.47	30.28
	SO_2 ($\mu g/m^3$)	21.26 ± 9.20	19.102	14.90 ± 4.43	14.506	19.21 ± 10.95	15.513
	CO (mg/m^3)	1.28 ± 0.56	1.0913	0.84 ± 0.18	0.8135	1.07 ± 0.46	0.9246
	O_3 ($\mu g/m^3$)	48.65 ± 32.18	38.486	49.94 ± 22.38	46.79	46.30 ± 34.30	39.032
Phase III	**PM_{10} ($\mu g/m^3$)**	292.28 ± 135.65	274.53	122.43 ± 52.57	116.78	145 ± 62.00	134.5
	$PM_{2.5}$ ($\mu g/m^3$)	98.22 ± 48.61	85.34	55.39 ± 28.96	50.67	58.09 ± 11.54	54.49
	NO_2 ($\mu g/m^3$)	52.02 ± 23.95	44.8	20.9 ± 7.46	18.26	26.72 ± 23.96	23.45
	SO_2 ($\mu g/m^3$)	20.75 ± 10.99	17.789	14.44 ± 4.13	13.365	12.61 ± 3.43	12.103
	CO (mg/m^3)	1.36 ± 0.68	1.0828	0.86 ± 0.17	0.8389	1.03 ± 0.26	0.9668
	O_3 ($\mu g/m^3$)	45.38 ± 32.34	33.598	62.31 ± 28.27	60.16	49.49 ± 28.39	40.44
Phase IV	**PM_{10} ($\mu g/m^3$)**	185.67 ± 77.79	173.51	157.4 ± 76.17	160	131.02 ± 147.26	99.18
	$PM_{2.5}$ ($\mu g/m^3$)	68.7 ± 30.24	65.4	58.82 ± 39.20	43.982	45.88 ± 46.70	34.257
	NO_2 ($\mu g/m^3$)	51.06 ± 22.73	46.93	29.94 ± 15.19	23.59	25.79 ± 12.40	21.843
	SO_2 ($\mu g/m^3$)	18.25 ± 7.72	16.729	16.69 ± 8.11	14.03	9.08 ± 2.00	8.467
	CO (mg/m^3)	1.261 ± 0.50	1.1482	1.01 ± 0.32	0.9172	0.87 ± 0.30	0.7859
	O_3 ($\mu g/m^3$)	47.33 ± 36.51	32.104	54.31 ± 29.16	50.363	38.21 ± 22.49	32.231

sectors, such as weekend curfew restriction and night curfew for movement of people, played a significant role. Less contribution from different sources had a positive contribution to improving the air quality (Nagpure et al. 2013). The phase-wise changes occurred and its hourly mean average is depicted in Table 2. The $PM_{2.5}$ concentration during the monitoring period shows a similar pattern as of PM_{10}, and it can be seen in Figs. 5, 6, and 7. The gaseous parameters showed different patterns influenced by various primary and secondary factors, such as the ozone concentration showed

an inverse pattern during the year 2020. It may be due to the photochemical reactions of NOx and VOCs influenced by the gases present in the atmosphere with sunlight (Gautam et al. 2021a; Naqvi et al. 2021).

Box plot model analysis

The study analyzed the variations in PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 concentrations during the different phases from 2019 to 2021. The averaged maximum and minimum concentrations show the range in which the particles are found during those different phases. The pictorial representation of the box plot analysis is depicted in Fig. 8. The maximum and minimum concentration of CO during 2019 in different lockdown phases was 3.3 mg/m^3 and 0.5 mg/m^3 (Phase I), 3.6 and 0.6 mg/m^3 (Phase II), 3.9 mg/m^3 and 0.4 mg/m^3 (Phase III), and 3.3 mg/m^3 and 0.7 mg/m^3 (Phase IV) with mean value of 1.2 ± 0.5 mg/m^3, 1.2 ± 0.5 mg/m^3, 1.3 ± 0.6 mg/m^3, and 1.2 ± 0.5 mg/m^3. Similarly, phase-wise values of O_3 for 2020 and 2021 can be seen in Fig. 8 and Table 2. During that period, the maximum and minimum concentrations were 95.9 µg/m^3 and 7.5 µg/m^3. The phase and year-wise concentration variation can be seen in Fig. 8 and Table 2. In 2020, ozone concentration was observed to be high due to secondary formation (Naqvi et al. 2021; Gautam et al. 2021b).

When SO_2 was considered, it showed a decrease in concentration in 2020 compared to 2019 (Fig. 4), but comparing 2020 and 2021, it showed an increment in two phases, P-I and P-II, and reduction during P-III and P-IV in 2021 (Fig. 3). The maximum reduction percentage was 45.5% in P-IV 2021 with BY 2019, whereas the increment was 42.9% in P-I 2021 with BY 2020 (Table 1). The recorded minimum concentration was 6.2 µg/m^3 during P-IV of 2021 and the maximum was 69.7 µg/m^3 during P-III of 2019.

Pearson correlation coefficient analysis

Figure 9 represents the Pearson correlation (PC) coefficients with a range of − 100 to 100. The value 100 represents "perfect correlation," 0 represents "no correlation," and − 100 represents "perfect inverse correlation." Pearson correlation analysis was done for particulate matter ($PM_{2.5}$, PM_{10}), gaseous pollutants (NO_2, SO_2, CO, and O_3) and meteorological parameters (*WS* wind speed, *WD* wind direction, *SR* solar radiation, *RH* relative humidity, *AT* ambient temperature). High correlations were found in O_3 and AT, with values ranging from 71 to 87, which shows evidence of the secondary formation of O_3 with respect to temperature. The correlation value between $PM_{2.5}$ and SO_2 falls in the range of 69–89. Similarly, NO_2 with CO showed a value in the range of 79–92, while an inverse correlation was found between NO_2 and WS (− 53 to − 75). The correlation was also found

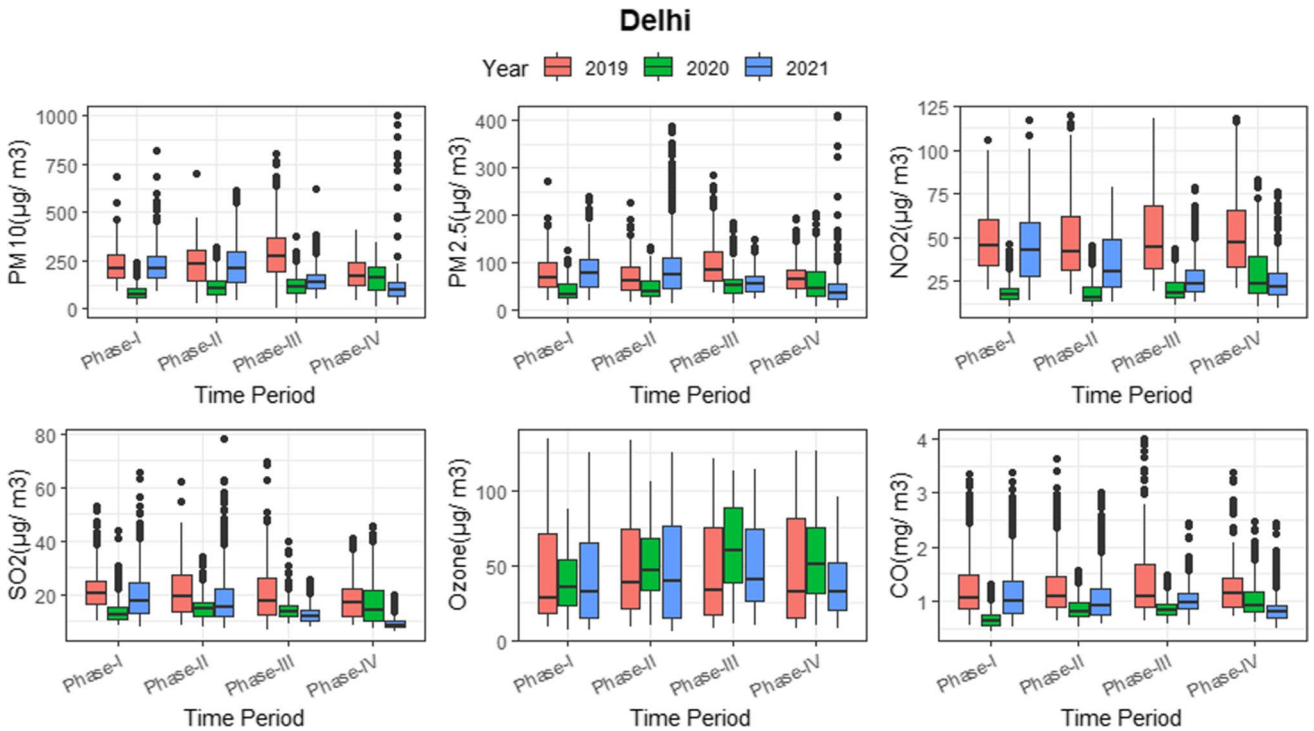


Fig. 8 Box plot variation of concentration of pollutants from 2019 to 2021

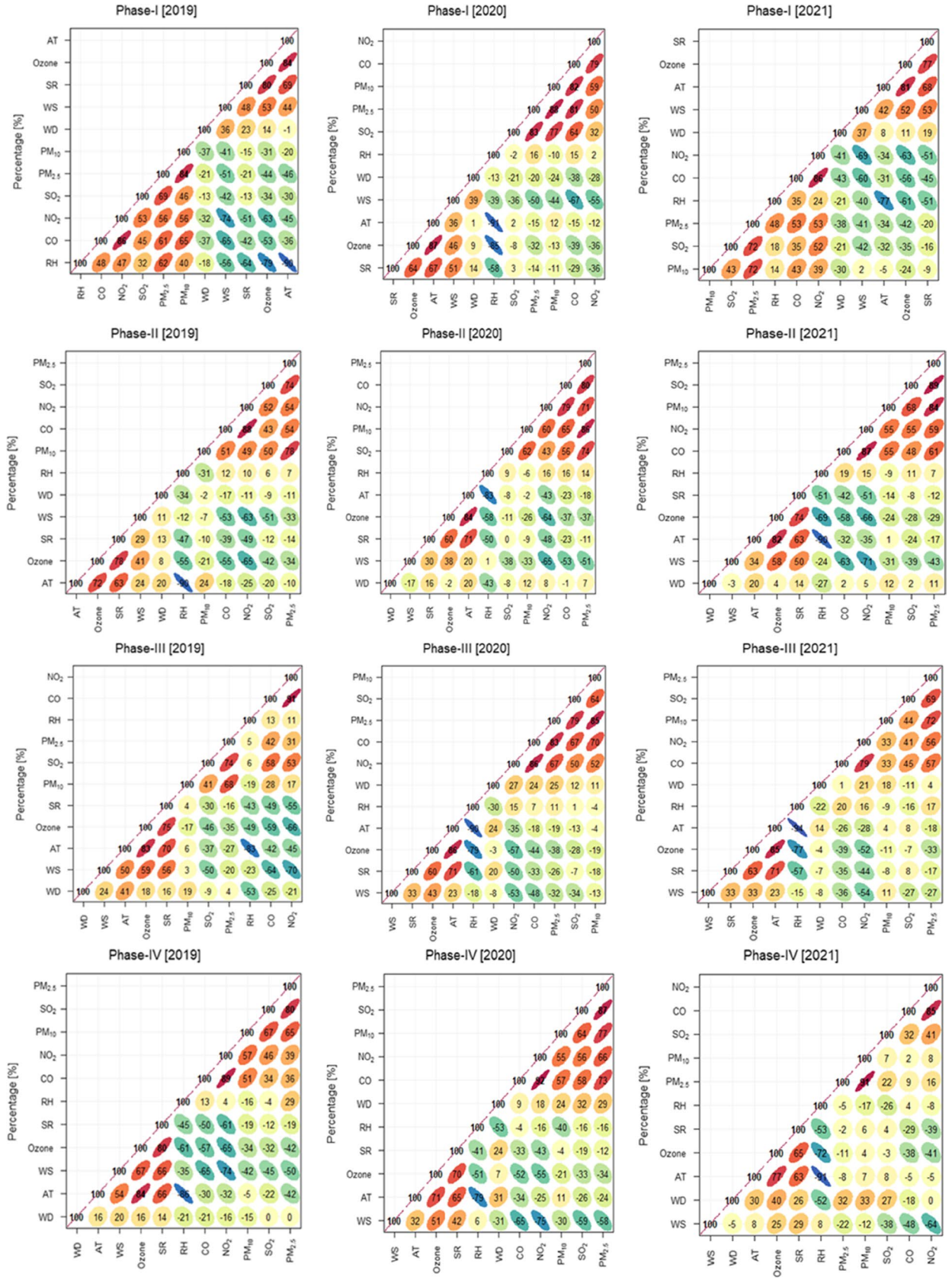


Fig. 9 Correlation matrix showing the relationships between all the parameters phase-wise during 2019–2021

varying year to year and phase to phase. The results show a high homogeneity between different phases and years.

Bivariate polar plot analysis

Bivariate polar plots examine the role of WS and WD on pollution concentrations. In Figs. 10 and 11, the WS and WD are analyzed for $PM_{2.5}$ and O_3. The pattern of the polar plot was quite similar for all parameters. Few other plots can also be referred in the supplementary section (Figures S4–S6.). The lower $PM_{2.5}$ concentrations occur under WS > 3 m/s and show small directional dependence. Similarly, high concentrations of $PM_{2.5}$ under WS < 2 m/s show that the local sources dominate $PM_{2.5}$ concentrations. Resuspension of dust may be the possible reason for this. Those local sources include road transport emissions and domestic heating. Low WS helps accumulate $PM_{2.5}$ emitted by local sources near the monitoring location, resulting in high $PM_{2.5}$ concentrations. At higher WS, $PM_{2.5}$ concentrations are recorded low with possible dispersion towards the WD. The bivariate polar analysis for O_3 shows low concentration at lower WS < 1 m/s and high concentration when WS is

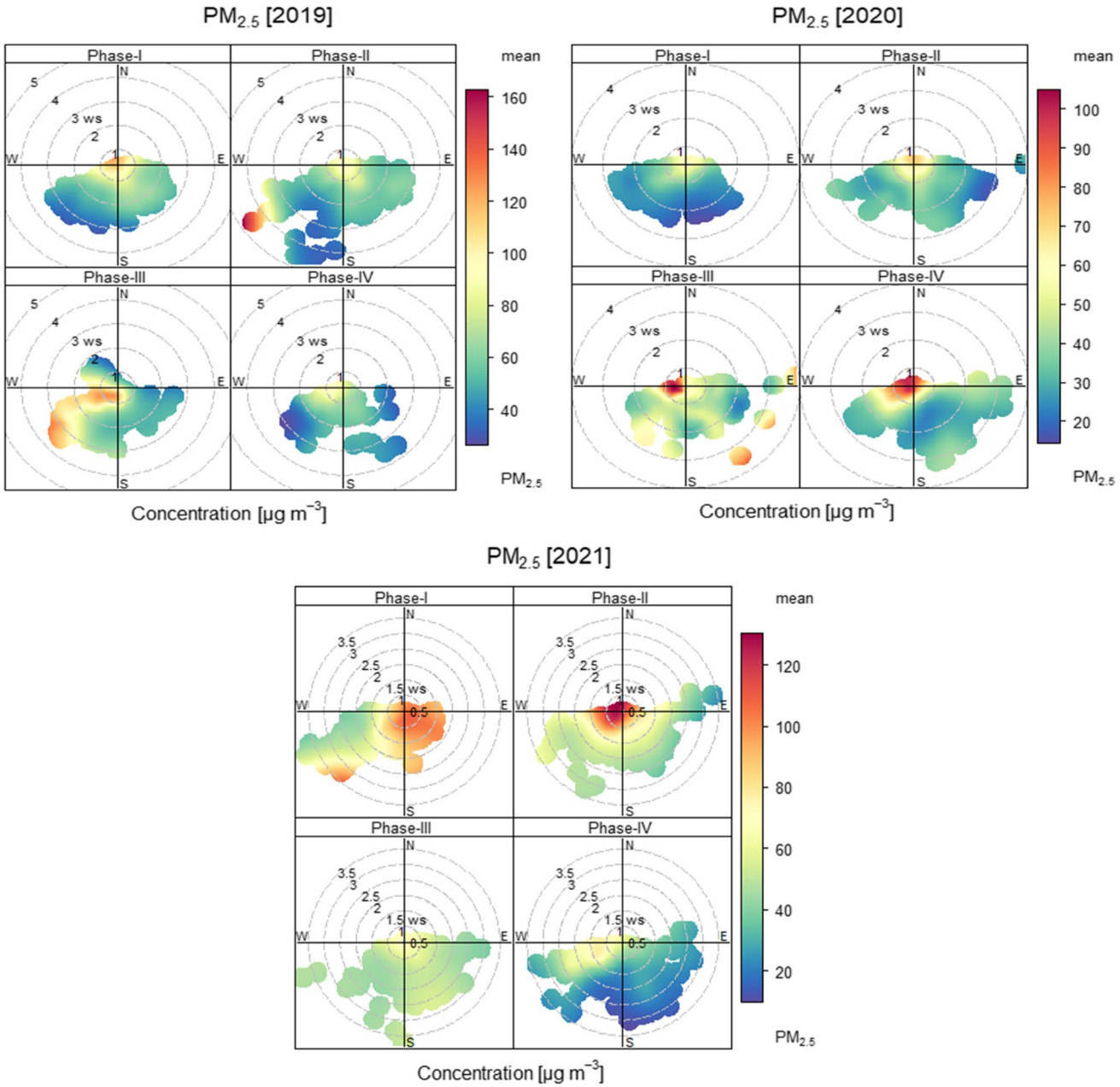


Fig. 10 Bivariate polar plots of $PM_{2.5}$ in Phases I, II, III, and IV for 2019–2021

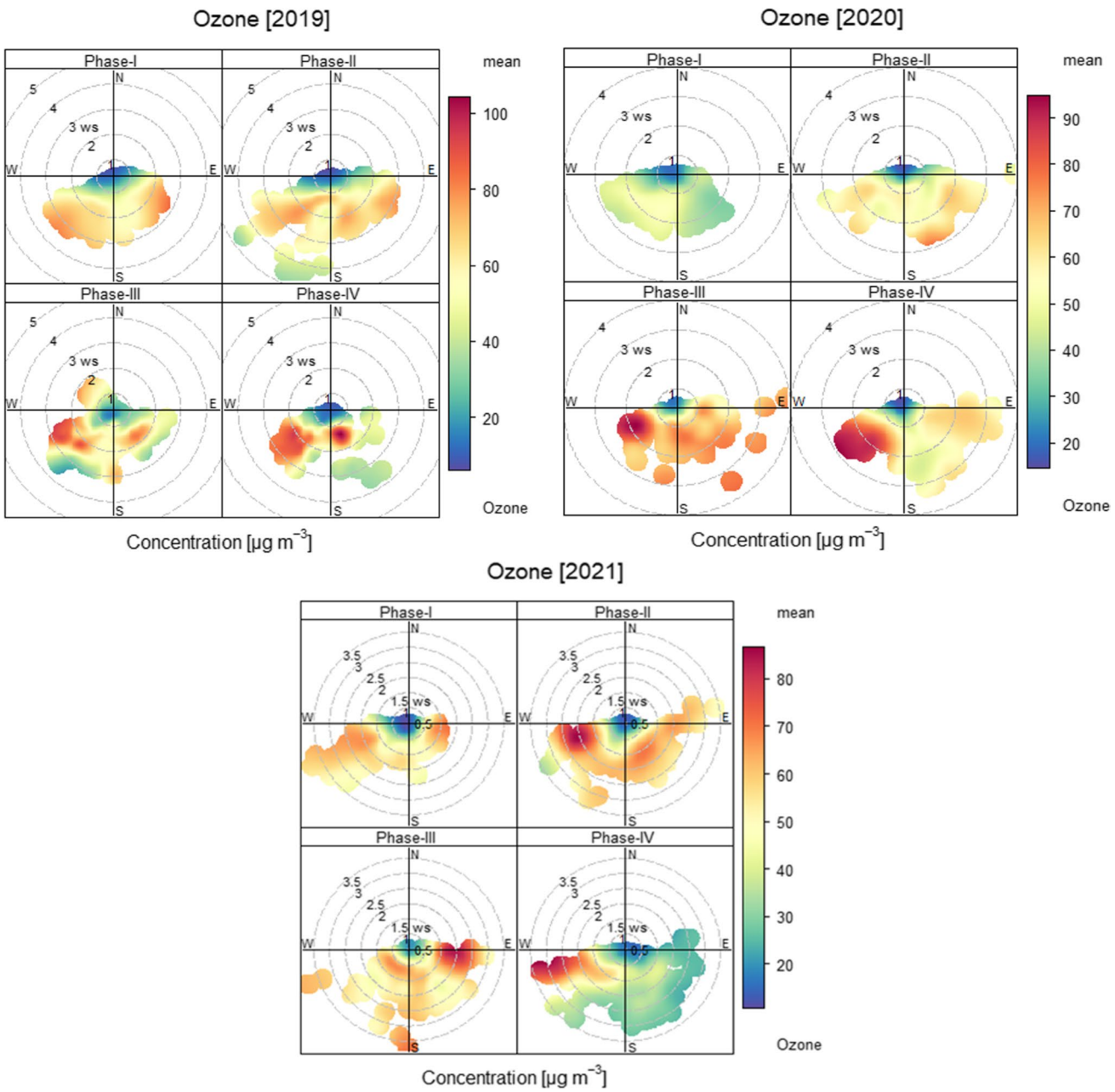


Fig. 11 Bivariate polar plots of O_3 in Phases I, II, III, and IV for 2019–2021

above 1.5 m/s. The previous studies revealed that PM_{10} is dominated by the construction sources and dust emissions in Delhi (Hazarika et al. 2016; Pant et al. 2015). A recent study reported that around 79% of PM_{10} originates from sources such as road dust resuspension (Nagpure et al. 2016; Singh et al. 2019). The boundary layer conditions, meteorology are also playing a significant role in the dispersion of the pollutants of that particular city.

Conclusion

The pandemic had influenced to switch off the anthropogenic sources globally in 2020. In India, the second pandemic wave created enormous pressure on government and healthcare systems in 2021, especially in Delhi. Even though it caused many tragedies in the economy and life of people, it showed fruitful results in pollution reduction. Many Indian cities experienced a reduction in air

pollution, including Delhi. After the era of urbanization, very first time, the air quality in Delhi dropped down dramatically. The overall $PM_{2.5}$ concentration reduction was found to be 35% during 2020. But at the same time, an increase in O_3 concentration of about 10% was observed in all the phases. It may be due to the secondary formation of O_3 by solar radiation and other VOCs. This shows that O_3 is not only associated with direct emissions in the city. The percentage change in P-I for parameters PM_{10}, $PM_{2.5}$, NO_2, SO_2, CO, and O_3 is 167%, 104.3%, 142.3%, 42.9%, 70%, and 6.8% in 2021, respectively. Similarly for P-II, it is 109.9%, 106.6%, 92.3%, 28.9%, 26.5%, and − 7.2%. The concentration in 2020 was recorded minimally and can act as baseline data. The study revealed that the concentration in 2020 was reduced multiple folds compared to 2019. The average mean concentration in 2019 was 226.4 ± 81.8 µg/m^3 (P-I), 227.5 ± 101.2 µg/m^3 (P-II), 292.2 ± 135.6 29 µg/m^3 (P-III), and 185.6 ± 77.7 µg/m^3 (P-IV), respectively. It reduced to 84.1 ± 38.5 µg/m^3, 112.4 ± 52.3 µg/m^3, 122.4 ± 52.5 µg/m^3, and 157.4 ± 76.1 µg/m^3 in 2020. The lower $PM_{2.5}$ concentrations occur under WS > 3 m/s and high concentrations during WS < 2 m/s. While comparing 2020 and 2021, the concentration was slightly increased in 2021, but the concentration in 2021 was still lower than that of 2019. High correlations were found in O_3 and AT, with values ranging from 71 to 87, which shows evidence of the secondary formation of O_3 with respect to temperature. The correlation value between $PM_{2.5}$ and SO_2 falls in the range of 69–89. It shows that the air quality in Delhi is improved to better standards in the year 2020 and 2021. Anthropogenic source's contribution plays a crucial role in influencing the pattern of air quality. The government of Delhi has implemented many amendments (e.g., ban on diesel generators, ban on crackers usage, odd–even scheme) to reduce air pollution from time to time when it reaches very high beyond the standards. The situation created by the pandemic paved the way for improvement in various sustainable approaches. The study outcome can be used to develop the planned switch-off of anthropogenic sources in the future to improve air quality.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12517-022-10567-8.

Acknowledgements The authors are thankful to the Central Pollution Control Board (CPCB) for providing a platform for accessing data from various locations and also other agencies such as the Delhi Pollution Control Committee (DPCC), India Meteorological Department (IMD), and Indian Institute of Tropical Meteorology (IITM) for providing data facility and members of the Advanced Air and Acoustics Research Laboratory, Department of Environmental Engineering, Delhi Technological University (DTU), for their support to conduct this study.

Author contribution Vignesh Mohan—Conceptualization, visualization, formal analysis, reviewing, and editing.

Rajeev Kumar Mishra—Investigation, discussion, reviewing, and editing.

Data availability Data are taken from public domain resources (CPCB website).

Code availability Not applicable.

Declarations

Competing interests The authors declare no competing interests.

References

Agarwal A, Kaushik A, Kumar S, Mishra RK (2020) Comparative study on air quality status in Indian and Chinese cities before and during the COVID-19 lockdown period. Air Qual Atmos Health 13:1167–1178. https://doi.org/10.1007/s11869-020-00881-z

Bera B, Bhattacharjee S, Shit PK, Sengupta N, Saha S (2021) Significant impacts of COVID 19 lockdown on urban air pollution in Kolkata (India) and amelioration of environmental health. Environ Dev Sustain 23:6913–6940. https://doi.org/10.1007/s10668-020-00898-5

Bherwani H, Nair M, Musugu K, Gautam S, Gupta A, Kapley A, Kumar R (2020) Valuation of air pollution externalities: comparative assessment of economic damage and emission reduction under COVID-19 lockdown. Air Qual Atmos Health 13:683–694. https://doi.org/10.1007/s11869-020-00845-3

Das M, Das A, Sarkar R, Saha S, Mandal P (2021) Regional scenario of air pollution in lockdown due to COVID-19 pandemic: Evidence from major urban agglomerations of India. Urban Clim 100821. https://doi.org/10.1016/j.uclim.2021.100821

Dhaka SK, Chetna, Kumar V, Anwar VP, Dimri AP, Singh N, Patra PK, Matsumi Y, Akigawa MT, Akayama TN, Yamaji K, Kajino M, Misra P, Hayashida S (2020) PM2.5 diminution and haze events over Delhi during the COVID 19 lockdown period: an interplay between the baseline pollution and meteorology. Sci Rep 10:13442. https://doi.org/10.1038/s41598-020-70179-8

Hamaa SML, Kumara P, Harrison RM, Bloss WJ, Kharee M, Mishra S, Namdeo A, Sokhi R, Goodman P, Sharma C (2020) Four-year assessment of ambient particulate matter and trace gases in the Delhi-NCR region of India. Sustain Cities Soc pp 102003. https://doi.org/10.1016/j.scs.2019.102003

Hazarika N, Srivastava A, Das A (2016) Quantification of particle bound metallic load and PAHs in urban environment of Delhi, India: Source and toxicity assessment. Sustain Cities Soc. https://doi.org/10.1016/j.scs.2016.11.010

He G, Pan Y, Tanaka T (2020) The short-term impacts of COVID-19 lockdown on urban air pollution in China. Nature Sustain vol 3:1005–1011.

Higham JE, Ram´ırez CA, Green MA, Morse AP (2021) UK COVID-19 lockdown: 100 days of air pollution reduction? Air Qual Atmos Health 14:325–332. https://doi.org/10.1007/s11869-020-00937-0

Gautam AS, Dilwaliya NK, Srivastava A, Kumar S, Bauddh K, Siingh D, Shah MA, Singh K, Gautam S (2021a) Temporary reduction in air pollution due to anthropogenic activity switch off during COVID 19 lockdown in northern parts of India. Environ Dev Sustain 23:8774–8797. https://doi.org/10.1007/s10668-020-00994-6

Gautam AS, Kumar S, Gautam S, Anand A, Kumar R, Joshi A, Bauddh K, Singh K (2021b) Pandemic induced lockdown as a boon to the Environment: trends in air pollution concentration

across India. Asia-Pacific J Atmos Sci. https://doi.org/10.1007/s13143-021-00232-7

Karuppasamy MB, Seshachalam S, Natesan U, Ayyamperumal R, Karuppannan S, Gopalakrishnan G, Nazir N (2020) Air pollution improvement and mortality rate during COVID-19 pandemic in India: global intersectional study. Air Qual Atmos Health 13:1375–1384. https://doi.org/10.1007/s11869-020-00892-w

Khaiwal R, Singh T, Biswal A, Singh V, Mor (2021) Impact of COVID-19 lockdown on ambient air quality in megacities of India and implication for air pollution control strategies. Environ Sci Pollut Res 28:21621–21632. https://doi.org/10.1007/s11356-020-11808-7

Kumar P, Hama S, Omidvarborna H, Sharma A, Sahani J, Abhijith KV, Debele SE, Zavala-Reyes JC, Barwise Y, Tiwari A (2020) Temporary reduction in fine particulate matter due to 'anthropogenic emissions switch-off' during COVID-19 lockdown in Indian cities. Sustain Cities Soc 102382. https://doi.org/10.1016/j.scs.2020.102382

Kumar R, Mishra RK, Chandra S, Hussain A (2021) Evaluation of urban transport-environment sustainable indicators during Odd-Even scheme in india. Environ Dev Sustain. https://doi.org/10.1007/s10668-021-01353-9

Kumari P, Toshniwal D (2020) Impact of lockdown measures during COVID-19 on air quality– A case study of India. Int J Environ Health Res. https://doi.org/10.1080/09603123.2020.1778646

Li L, Li Q, Huang L, Wang Q, Zhu A, Xua J, Liu Z, Li H, Shi L, Li R, Azari M, Wang Y, Zhang X, Liu Z, Zhu Y, Zhang K, Xue S, Ooi MCG, Zhang D, Chan A (2020) Air quality changes during the COVID-19 lockdown over the Yangtze River Delta Region: An insight into the impact of human activity pattern changes on air pollution variation. Sci Total Environ 139282. https://doi.org/10.1016/j.scitotenv.2020.139282

Mahato S, Pal S, Ghosh KG (2020) Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India. Sci Total Environ 139086. https://doi.org/10.1016/j.scitotenv.2020.139086

Mishra RK, Pandey A, Pandey G, Kumar A (2019) The effect of odd-even driving scheme on PM2.5 and PM1.0 emission. Trans Res Part D 541–552. https://doi.org/10.1016/j.trd.2019.01.005

Nagpure AS, Sharma K, Gurjar BR (2013) Traffic induced emission estimates and trends (2000–2005) in megacity Delhi. Urban Clim pp 61–73. https://doi.org/10.1016/j.uclim.2013.04.005

Nagpure AS, Gurjar BR, Kumar V, Kumar P (2016) Estimation of exhaust and non-exhaust gaseous, particulate matter and air toxics emissions from on-road vehicles in Delhi. Atmos Environ 118–124. https://doi.org/10.1016/j.atmosenv.2015.12.026

Naqvi HR, Datta M, Mutreja G, Siddiqui MA, Naqvi DF, Naqvi AR (2021) Improved air quality and associated mortalities in India under COVID-19 lockdown. Environ Pollut 115691. https://doi.org/10.1016/j.envpol.2020.115691

Pant P, Shukla A, Kohl SD, Chow JC, Watson JG, Harrison RM(2015) Characterization of Ambient PM2.5 at a Pollution Hotspot in New Delhi, India and Inference of Sources. Atmos Environ.

Sarfraz M, Shehzad K, Shah SGM (2020) The impact of COVID-19 as a necessary evil on air pollution in India during the lockdown. Environ Pollut. https://doi.org/10.1016/j.envpol.2020.115080

Sikarwar A, Rani R (2020) Assessing the immediate effect of COVID-19 lockdown on air quality: A case study of Delhi, India. Res Square. https://doi.org/10.21203/rs.3.rs-31822/v1

Singh V, Biswal A, Kesarkar AP, Mor S, Ravendra K (2019) High resolution vehicular PM10 emissions over megacity Delhi: Relative contributions of exhaust and non-exhaust sources. Sci Total Environ 0048–9697. https://doi.org/10.1016/j.scitotenv.2019.134273

Singh V, Singh S, Biswal A, Kesarkar AP, Mor S, Ravindra K (2020) Diurnal and temporal changes in air pollution during COVID-19 strict lockdown over different regions of India. Environ Pollut 115368. https://doi.org/10.1016/j.envpol.2020.115368

Srivastava S, Kumar A, Bauddh D, Gautam AS, Kumar S (2020) 21 Day Lockdown in India Dramatically Reduced Air Pollution Indices in Lucknow and New Delhi, India. Bull Environ Contam Toxicol 105:9–17. https://doi.org/10.1007/s00128-020-02895-w

United Nations, Department of Economic and Social Affairs, Population Division (2019) World UrbanizationProspects: The 2018 Revision (ST/ESA/SER.A/420). New York: United Nations. https://population.un.org/wup/publications/Files/WUP2018-Report.pdf

Vadrevu KP, E aturu A, Biswas S, Lasko K, Sahu S, Garg JK, Justice C (2020) Spatial and temporal variations of air pollution over 41 cities of India during the COVID 19 lockdown period. Sci Rep 10:16574 https://doi.org/10.1038/s41598-020-72271-5

Venter ZS, Aunan K, Chowdhury S, Lelievel J (2020) COVID-19 lockdowns cause global air pollution declines. https://doi.org/10.1073/pnas.2006853117

WORLD AIR QUALITY REPORT (2019) https://www.iqair.com/world-most-polluted-cities/world-air-quality-report-2019-en.pdf

The work has not been published anywhere previously in any form and it is not in consideration for publication elsewhere. The publication is approved by all authors.

Evolving Systems
https://doi.org/10.1007/s12530-022-09458-w

ORIGINAL PAPER

A player unknown's battlegrounds ranking based optimization technique for power system optimization problem

Kapil Deo Bodha[1] · **V. Mukherjee**[1] · **Vinod Kumar Yadav**[2]

Received: 20 July 2021 / Accepted: 28 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

A player unknown's battlegrounds (PUBG) ranking based optimization (PRBO) algorithm is proposed. It is based on the ranking scheme of a popular battle royale game PUBG where each team tries to win. The performance of a team is based on its rank. Ranks are represented through six different equations. Six mathematical benchmark functions and five computationally expensive CEC 2015 functions are used to test the performance of the proposed algorithm. The obtained results are compared with various state-of-the-art well-known metaheuristic optimization algorithms. The proposed PRBO is also employed for power system dispatch problem for 3-unit, 5-unit and a 26-unit system. Additionally, it is also used to formulate the combined economic emission dispatch problem of IEEE 30-bus, 6-unit system for various loads. Emissions of SO_2, CO_2 and NO_x are also considered. A fuzzy concept is adopted to find out the best compromise solution among the Pareto optimal solutions. The simulation results indicate that the presented PRBO algorithm performs satisfactorily for all the tested instances, and it can obtain the optimal solution with faster convergence mobility.

Keywords Combined economic emission dispatch · Cubic criterion function · Emission reduction · Global optimum solution · Greenhouse gases · Player unknown's battlegrounds (PUBG)

Abbreviations

ABC	Artificial bee colony
BBO	Biogeography-based optimization
BWO	Black widow optimization
BCS	Best compromise solution
CAEO	Chaotic artificial ecosystem-based optimization algorithm
CCF	Cubic criterion function
CCP	Carbon capture plants
CEED	Combined EED
CEP	Classical evolutionary programming
CSA	Crow search algorithm
CSAISA	Chaotic self-adaptive interior search algorithm
CTLBO	Chaotic TLBO
DEED	Dynamic EED
ED	Economic dispatch
EED	Economic emission dispatch
FA	Firefly algorithm
FMOPSO	Fuzzified multi-objective PSO
GA	Genetic algorithm
GSA	Gravitational search algorithm
GSAPSO	Hybrid gravitational search PSO
HOA	Hurricane optimization algorithm
ISA	Interior search algorithm
KSO	Kernel search algorithm
LR	Lagrange relaxation
LSM	Least square method
MPSO_SSM	Improved PSO integrated with a simplex search method
MRFO-GBO	Manta ray foraging optimization integrated with gradient-based optimizer
NSGA-II	Nondominated sorting GA II

✉ Kapil Deo Bodha
bodhas@gmail.com

[1] Department of Electrical Engineering, Indian Institute of Technology, Dhanbad, India

[2] Department of Electrical Engineering, Delhi Technological University, Delhi, India

Published online: 14 August 2022

PUBG	Player unknown's battlegrounds
PSO	Particle swarm optimization
QBAT	Quantum behaved bat algorithm
QPSO	Quantum inspired PSO
QO-TLBO	Quasi-oppositional TLBO
QITFA	Quantum inspired TFA
QWOA	Quantum WOA
SA	Simulated annealing
SCA	Sine cosine algorithm
SOS	Symbiotic organisms search algorithm
SDSM	Simplified direct search method
SQP	Sequential quadratic programming
TFA	Tidal FA
TLBO	Teaching learning based optimization
VPL	Valve point loading
WOA	Whale optimization algorithm
CO_2	Carbon dioxide
$E(P_i)$	Emission objective function of the *i*th generating unit
fn_j^{max} and fn_j^{min}	Maximum and minimum values of the *i*th objective function
$J_{i,CO_2}(P_i)$	Emission objective function of CO_2
$J_{i,NO_x}(P_i)$	Emission objective function of NOx
$J_{i,SO_2}(P_i)$	Emission objective function of SO_2
$K_{i(max/max)}$	Max-max penalty factor of the *i*th generating unit
n	Total numbers of generators
NO_x	Nitrogen oxide
P	Total power generated
P_D	Demanded power
P_L	Transmission line losses
P_{slack}	Power of the slack generator
$P_{i,min}$	Minimum generating limit of the *i*th generating unit
$P_{i,max}$	Maximum generating limit of the *i*th generating unit
S	Population size/Number of team
SO_2	Sulphur dioxide
t_{max}	Maximum number of iterations
$X_{conquerornew,p}$	Updated position/Number of kills by the *p*th player of the conqueror team
$X_{average,p}$	Average number of kills by the *p*th player of all the teams
$X_{crown,p}$	Number of kills by the *p*th player of the crown ranked team
$X_{ace,p}$	Number of kills by the *p*th player of the ace ranked team.
$X_{diamond,p}$	Number of kills by the *p*th player of the diamond ranked team
$X_{platinum,p}$	Number of kills by the *p*th player of the platinum ranked team
$X_{rookie,p}$	Number of kills by the *p*th player of the rookie ranked team
$X_{i,p}$	Number of kills of randomly selected *p*th player of the *i*th team
$\alpha_{0(CO_2),i}, \alpha_{1(CO_2),i}, \alpha_{2(CO_2),i}, \alpha_{3(CO_2),i}$	Emission coefficients of CO_2 for the *i*th generating unit
$\alpha_{0(NO_x),i}, \alpha_{1(NO_x),i}, \alpha_{2(NO_x),i}, \alpha_{3(NO_x),i}$	Emission coefficient of NO_x for the *i*th generating unit
$\alpha_{0(SO_2),i}, \alpha_{1(SO_2),i}, \alpha_{2(SO_2),i}, \alpha_{3(SO_2),i}$	Emission coefficient of SO_2 for the *i*th generating unit
$\omega_{0,i}, \omega_{1,i}, \omega_{2,i}, \omega_{3,i}$	Fuel cost coefficient of the *i*th generating unit
γ	Membership function

1 Introduction

Classical optimization methods like lambda-iteration, Newton–Raphson, Lagrange relaxation, weighted min–max and quadratic programming may not produce the best answer for complex engineering situations. As a result, in recent decades, numerous metaheuristic optimization algorithms have been seen to handle various complex problems in diverse domains of applications such as science, engineering, management and economics. Many metaheuristic optimization techniques have been proposed by the scholars in last few decades. To determine best results, each of these algorithms employs a distinct methodology.

Combined economic emission dispatch (EED) (CEED) is one such complex optimization problem of electricity utility

companies. It refers to scheduling of generators in such a way to minimize the fuel cost as well the emission without violating any constraint. It is a highly complex optimization problem, as the search space is very limited owing to the number of constraints. The economic dispatch (ED) refers to scheduling of generators in a way to minimize the fuel cost without violating any constraints. To tackle the CEED/ED problem of power system, many optimization strategies have been proposed in recent state-of-the-art literatures.

In one of the earliest work related to CEED, the weighted min–max technique along with fuzzy set theory was used to resolve the CEED problem of a 6-unit system (Dhillon et al. 1993). In (Jeyakumar et al. 2006), particle swarm optimization (PSO) technique was used to formulate the CEED problem of different test systems and obtained results were then collated with classical evolutionary programming (CEP) technique to prove the superiority of the PSO technique. Spiral optimization algorithm was proposed in (Benasla et al. 2014) for the solution of the CEED problem. It was applied on 3-unit, 6-unit and 40-unit system and the results were compared with other state-of-the-art optimization techniques. CEED and ED problems were formulated using flower pollination algorithm in (Abdelaziz et al. 2016). The proposed technique was effective even for large systems. The mating behavior of fireflies was replicated to develop the firefly algorithm (FA) which was then used to address the CEED problem of power systems. The results were compared with PSO technique. The FA provided marginal improvement over PSO. It has good convergence characteristic but requires adjusted or tuned values of its many parameters to get efficient operation (see (Apostolopoulos and Vlachos 2011)).

A crow search algorithm (CSA), based on food searching behavior of crows, was utilized for parameter estimation of constants used in ED equation (see (Askarzadeh and Gharibi 2018)). The simulation results showed that the proposed technique is better than the least square method (LSM), PSO, genetic algorithm (GA) and artificial bee colony techniques. The optimal tuning of interval type-2 fuzzy controllers is achieved using a slime mould algorithm (SMA) and a SMA with information feedback model F1 (SMA-F1). SMA and SMA-F1 outperform other metaheuristic algorithms, as demonstrated by the results presented in (Precup et al. 2021). To represent the automobile motion and any detected disruptions, a half car model was developed and applied. The measuring inputs are front and rear left suspension reflection, vertical acceleration, and roll rate. To categorize the discovered abnormalities, threshold-based detection approaches were applied. The system obtained an 86% labeling accuracy and a 2% false negative rate (FNR) (Li et al. 2017).

It was observed in (Hossein et al. 2020) that when PSO and GA algorithms are applied using neural network, they perform much better than their classical individual counterparts. A novel kernel search algorithm (KSO) was proposed to solve the ED problem considering valve point loading (see (Dong and Wang 2020)). The proposed algorithm was applied on IEEE 30-bus system, 10-unit system and 40-unit system with various load profiles. The results proved the supremacy of KSO over other algorithms. Gravitational search algorithm (GSA), based on force of attraction between heavenly bodies, was implemented to resolve the CEED problem in (Shaw et al. 2012). GSA suffers from slow convergence as there is no exchange of information between the particles. To address these issues, PSO is integrated with GSA (Mirjalili and Hashim 2010). A chaotic teaching–learning based optimization (TLBO) with levy flight (CTLBO) was applied to solve the economic load dispatch problem of various systems (refer (He et al. 2016)). The results were compared with basic TLBO, simulated annealing (SA) and GA technique. It was observed that CTLBO exhibited excellent convergence mobility and performed well compared to other techniques reported in the referred literature. In (Karthik et al. 2019), an interior search algorithm (ISA) was presented for scheduling of the power plant considering fuel cost and emission minimization. The technique was applied on 3-unit system, IEEE 30-bus system, 10-unit system, 20-unit system and Taiwan 40 generating unit system. The obtained results proved the dominance of ISA over other techniques.

In (Rajagopalan et al. 2019), a chaotic self-adaptive ISA (CSAISA) was used to handle the CEED issue while taking into account different security limitations and valve point effect. The results were compared with ISA, GA, PSO and differential evolution techniques. The CSAISA outperformed the other techniques in terms of speed and solution quality. A comparison between hurricane optimization algorithm (HOA) and sine cosine algorithm (SCA) was carried out in (El-Sehiemy et al. 2019). These two methods were used to counter the combined economic and ecological load dispatch

problems for several test systems. Based on the obtained results, it was concluded that HOA is a better alternative to SCA.

Mahdi et al. (2018) presented a quantum inspired PSO (QPSO) technique based on quantum mechanics principle. The proposed technique proved to perform better than SA, PSO and Lagrange relaxation techniques. The QPSO technique was further improved by Zhao et al. in (Xin-gang et al. 2020) and the same was applied to resolve economic and environmental dispatch of a microgrid.

CEED problem was framed using the cubic criteria function (CCF) in (Bodha et al. 2020) and a quantum inspired tidal firefly algorithm (QITFA) was used to optimize it. The QITFA fared well against quantum bat algorithm (QBA), QPSO, LR, PSO and SA. The performance of QITFA is dependent on the parameters which must be sensibly calibrated.

A quantum inspired time varying inertia weight GSA-PSO (QI-TVIW-GSAPSO) technique was employed to solve the dynamic economic emission dispatch (DEED) problem of a thermal solar hybrid power system. The DEED of a hybrid energy system is a mixed integer optimization problem. The QI-TVIW-GSAPSO effectively handles the complex optimization problem (Bodha et al. 2021).

Sakthivel et al. (2021) addressed the CEED problem for three test systems using a multi-objective squirrel search optimization approach. A modified squirrel search technique was also proposed which is based on the integration of squirrel search algorithm with Pareto dominance principle. The hybrid squirrel search algorithm performed much better than the squirrel search algorithm.

To address the EED problem of power plants, an algebraic (deterministic) method was used in (Barbosa-Ayala et al. 2020). The proposed technique gave better results than the sequential quadratic programming (SQP) algorithm and the nondominated sorting GA II (NSGA-II). EED problem with valve point loading of IEEE 30-bus system was solved using a modified Harris-hawk optimization approach in (Xia et al. 2021). The proposed approach utilizes the Pareto optimal solution to find the best optimum balanced solution. The technique performed better than the NSGA-II technique.

The constrained CEED problem was resolved using an equilibrium optimizer technique. This algorithm combines the dominant criteria for multi-objective functions that allow the decision-maker to detect all the Pareto boundaries of the constrained problem. The technique was successfully tested on a IEEE 30-bus system in (El-Shorbagy and Mousa 2021). To handle the CEED problem, an enhanced PSO coupled with a simplex search technique (MPSO SSM) was developed and reported in the literature. A price penalty factor (PPF) was employed to convert the multi-objective function into single-objective one. This eliminates the time intensive complications to deliver a workable generation schedule. The technique was compared with ten other optimization techniques in (Chopra et al. 2021). An improved PSO technique was proposed that utilizes a novel clone selection method to avoid stagnation of the PSO technique. It was used to manage the DEED problem in three test systems that had non-smooth fuel and emission functions (Qian et al. 2020). The presented results showed the supremacy of the proposed technique over the other algorithms. A weighting updated artificial bee colony technique was integrated with an environment-based demand response program to formulate the CEED problem. These findings were helpful in solving environmental problems for microgrid operators (Ryu and Kim 2020). In (Zhang et al. 2021), an improved symbiosis PSO algorithm with mutualistic symbiosis strategy was proposed. It uses three swarms to find the optimal solution. The proposed technique performed better than PSO, improved chaos PSO, and symbiotic organisms search (SOS) algorithm.

A chaotic artificial ecosystem-based optimization algorithm (CAEO) was proposed in (Hassan et al. 2021a). PPF was proposed to convert the four objective functions into single one. The CAEO algorithm was utilized for optimizing the economic load dispatch and three greenhouse gas emissions such as sulfur dioxide (SO_2), nitrous oxide (NO_x), and carbon dioxide (CO_2). The proposed CAEO technique performed better than the conventional artificial ecosystem-based optimization.

In this manuscript, a novel PRBO technique is applied to resolve the ED problem of 3-unit, 5-unit and 26-unit system and CEED problem of a 6-unit system. The technique is also validated on 6 benchmark functions and 5 computationally expensive CEC 2015 benchmark functions. The PRBO technique is based on ranking scheme of a popular battle royale game PUBG. Cubic criterion function (CCF) is used to formulate the dispatch problem, as cubic function represents the authentic response of a thermal generating unit more precisely than the classical quadratic functions (Mahdi et al. 2016). A max/max PPF is employed to convert the multi-objective function into single-objective function. A fuzzy set theory is adopted to find out the best compromise solution among the Pareto optimal solutions.

The main contributions of this paper are summarized in the following points:

a) A novel optimization technique based on PUBG ranking scheme is proposed. There are six different ranks, and the solutions are updated accordingly.
b) The performance of an optimization technique is greatly affected by the parameters chosen. The proposed PRBO technique does not have any parameters to set except defining the initial population size.
c) The constraint handling mechanism adopted in the dispatch problem has been clearly demonstrated that was missing from the literature.
d) Time complexity of the proposed metaheuristic technique is also presented.

The proposed technique is applied on 6 mathematical benchmark functions and 5 computationally expensive functions suggested in CEC 2015. It is also utilized to solve the ED and CEED problem of various power systems having varying degrees of complexities. The results are compared with the results surfaced in various state-of-the-art literatures such as LR, SA, PSO, QPSO, QITFA, QBA, SCA and MRFO-GBO techniques. The results prove that the proposed optimization technique is very effective in handling the optimization problem and can give the best results.

The rest of this paper is arranged as follows. In Sect. 2, the CEED problem formulation and constraint handling mechanism are presented. Section 3 describes the proposed PRBO algorithm. Section 4 reports results and discussion. The conclusions and scopes of future work of the paper are presented in Sect. 5.

2 Problem statements

2.1 Objective functions

There are two functions to be optimized viz., (a) fuel cost function and (b) emission function for three pollutants (like CO_2, NO_x and SO_2). The aim is to minimize both quantities simultaneously without violating any constraint. It can be written as in (1).

$$\text{Minimize } T(P) = [F(P), E(P)] \tag{1}$$

where $F(P)$ is the fuel cost (\$/h) to generate power P (MW) and $E(P)$ is the total pollutant emission produced (kg/h) in generating power P.

2.2 Factor of influence

Equation (1) summarizes the CEED problem. It depends on the variable P. It can be expressed as in (2)

$$P = [P_1, P_2, P_3 \ldots, P_n] \tag{2}$$

where powers generated by generating units are denoted by $P_1, P_2, P_3, \ldots, P_n$ and P_n is the power generated from the nth generating unit.

2.3 Fuel cost dispatch

The motive of ED is to optimize fuel cost to a minimum value without violating any constraint. It can be expressed as cubic function as shown in (3)

$$\text{Minimize } F(P) = \sum_{i=1}^{n} F_i(P_i) = \sum_{i=1}^{n} \left(\omega_{0,i} + \omega_{1,i}P_i + \omega_{2,i}P_i^2 + \omega_{3,i}P_i^3\right) \tag{3}$$

where $F_i(P_i)$ is the fuel cost of the ith generating unit in producing power P; $\omega_{0,i}, \omega_{1,i}, \omega_{2,i}$ and $\omega_{3,i}$ are the fuel cost coefficients of the ith generator.

2.4 Emission dispatch

Emission dispatching refers to reduction of the harmful pollutant emission as much as possible without violating any constraint. The total pollutant emission by the ith generating unit is denoted by $J_i(P_i)$ and the objective is to minimize the same (see (4))

$$\text{Minimize } E(P) = \sum_{i=1}^{n} J_i(P_i) = \sum_{i=1}^{n} \left(\alpha_{0,i} + \alpha_{1,i}P_i + \alpha_{2,i}P_i^2 + \alpha_{3,i}P_i^3\right) \tag{4}$$

where $\alpha_{0,i}, \alpha_{1,i}, \alpha_{2,i}$ and $\alpha_{3,i}$ are the emission coefficients of the ith generating unit.

2.5 Power balance constraint

The total output power of the generator should be equal to the sum of total demanded load and losses as given by (5)

$$\sum_{i=1}^{n_g} P_i = P_D + P_L \tag{5}$$

where P_i represents the power generated by the ith generator, P_D is total demanded power and P_L is the transmission line losses.

2.6 Power output limitations

The generating units are bounded by their minimum and maximum generating limits. These limits should not be violated while scheduling. It can be expressed as presented in (6)

$$P_{i,min} \le P_i \le P_{i,max} \tag{6}$$

where $P_{i,min}$ represents the lowest power generation limit of the ith generating unit and $P_{i,max}$ is the highest power generating capacity of the ith generator.

2.7 Handling equality constraints

To enforce the constraint given in (5), the slack bus generation method is employed (Radosavljević 2016). The pseudo code for constrained handling of the CEED is summarized in Algorithm 1.

2.8 Max/max PPF

The PPF is the ratio of fuel cost to emission value. There are basically three types of PPF such as (a) min–max PPF, (b) max-max PPF and (c) average PPF. The role of all penalty factors is to convert the multi-objective function into a single-objective function. This manuscript uses the max-max PPF, as it is a preferred choice for the CEED studies

Algorithm 1 The pseudo code for constraint handling mechanism in CEED

Input Fuel cost coefficients, emission coefficient, minimum power limit (P_{min}), maximum power limit (P_{max}), demanded power (P_d), number of generators (n), number of population or teams (S)

%Choose Generator-1 as slack generator and generate initial solutions for generators (2-*n*)

for i=1:S
  for j=2:n
    $P_{withoutslack}(i, j)= P_{min}(j)$+rand*($P_{max}(j)- P_{min}(j)$);
  end
end

% Generate solution for slack generator by using (5)

for i=1:S
  for j=2:n
    $P_{slack}(i,1)=P_d$- sum($P_{withoutslack}(i, j)$);
  end
end

% Checking limits of slack generator and imposing error, if out of bounds

for i=1:S
  if ($P_{slack}(i, 1)<P_{min}(1, 1)$)
    $error(i)= (P_{slack}(i, 1)-P_{min}(1, 1))^{\wedge}2$;
  elseif (Pslack(i, 1)>Pmax(1, 1))
    $error(i)$= (maxi(1, 1)- $P_{slack}(i, 1)$)$^{\wedge}2$;
  else
    $error(i)$=0
  end
  if $P_{slack}(i,1)=P_{min}(1, 1)$
    $error(i)$=0;
  elseif $P_{slack}(i,1)= P_{max}(1, 1)$
    $error(i)$=0;
  end
end

$a=P_{withoutslack}(:, 2: n)$;

P=[P_{slack}, a] **% Initial population generation of order [n×S]**

(Pitono 2014). The multi-objective function can be transformed into a single-objective function by utilizing max/max PPF ($K_{max/max}$) (as shown in (7)). The complete equation is shown in (8). The max/max PPF unit wise is expressed in (9)

$$\text{Minimize } T(P) = \left(F(P) + K_{max/max} * E(P)\right) = \sum_{i=1}^{n_g} F_i\left(P_i\right) + K_{i(max/max)} \sum_{i=1}^{n_g} E_i\left(P_i\right) \tag{7}$$

$$\text{Minimize } T(P) = \sum_{I=1}^{n_g} F_i\left(P_i\right) + K_{i\left(\frac{max}{max}\right),SO_2} J_{i,SO_2}\left(P_i\right) + K_{i\left(\frac{max}{max}\right),NO_x} J_{i,NO_x}\left(P_i\right) + K_{i\left(\frac{max}{max}\right),CO_2} J_{i,CO_2}\left(P_i\right) \tag{8}$$

where

$$K_{i\left(\frac{max}{max}\right)} = \sum_{i=1}^{n} \frac{F_i(P_{i,max})}{E_i\left(P_{i,max}\right)} \tag{9}$$

where $J_{i,SO_2}\left(P_i\right)$, $J_{i,NO_x}\left(P_i\right)$ and $J_{i,CO_2}\left(P_i\right)$ are the emission (kg/h) of SO_2, NO_x and CO_2, respectively. Cubic equations are used to express the emission pollutants equations (see (10, 11, 12))

$$J_{i,SO_2}\left(P_i\right) = \alpha_{0(SO_2),i} + \alpha_{1(SO_2),i} P_i + \alpha_{2(SO_2),i} P_i^2 + \alpha_{3(SO_2),i} P_i^3 \tag{10}$$

where $\alpha_{0(SO_2),i}, \alpha_{1(SO_2),i}, \alpha_{2(SO_2),i}$ and $\alpha_{3(SO_2),i}$ represent SO_2 emission coefficient of the ith generating unit.

$$J_{i,NO_x}\left(P_i\right) = \alpha_{0(NO_x),i} + \alpha_{1(NO_x),i} P_i + \alpha_{2(NO_x),i} P_i^2 + \alpha_{3(NO_x),i} P_i^3 \tag{11}$$

where $\alpha_{0(NO_x),i}, \alpha_{1(NO_x),i}, \alpha_{2(NO_x),i}$ and $\alpha_{3(NO_x),i}$ represent NO_x emission coefficient of the ith generating unit.

$$J_{i,CO_2}\left(P_i\right) = \alpha_{0(CO_2),i} + \alpha_{1(CO_2),i} P_i + \alpha_{2(CO_2),i} P_i^2 + \alpha_{3(CO_2),i} P_i^3 \tag{12}$$

where $\alpha_{0(CO_2),i}, \alpha_{1(CO_2),i}, \alpha_{2(CO_2),i}$ and $\alpha_{3(CO_2),i}$ represent CO_2 emission coefficient of the *i*th generating unit. The associated terms involved are expressed in (13)–(18) (Ziane et al. 2017).

$$K_{i\left(\frac{max}{max}\right),SO_2} = \sum_{i=1}^{n} \frac{F_i(P_{i,max})}{J_{i,SO_2}\left(P_{i,max}\right)} \tag{13}$$

$$K_{i\left(\frac{max}{max}\right),SO_2} = \sum_{i=1}^{n} \frac{\omega_{0,i} + \omega_{1,i} P_{i,max} + \omega_{2,i} P_{i,max}^2 + \omega_{3,i} P_{i,max}^3}{\alpha_{0(SO_2),i} + \alpha_{1(SO_2),i} P_{i,mx} + \alpha_{2(SO_2),i} P_{i,max}^2 + \alpha_{3(SO_2),i} P_{i,max}^3} \tag{14}$$

$$K_{i\left(\frac{max}{max}\right),NO_x} = \sum_{i=1}^{n} \frac{F_i(P_{i,max})}{J_{i,NO_x}\left(P_{i,max}\right)} \tag{15}$$

$$K_{i\left(\frac{max}{max}\right),NO_x} = \sum_{i=1}^{n} \frac{\omega_{0,i} + \omega_{1,i} K_{i,max} + \omega_{2,i} K_{i,max}^2 + \omega_{3,i} K_{i,max}^3}{\alpha_{0(NO_x),i} + \alpha_{1(NO_x),i} P_{i,max} + \alpha_{2(NO_x),i} P_{i,max}^2 + \alpha_{3(NO_x),i} P_{i,max}^3} \tag{16}$$

$$K_{i\left(\frac{max}{max}\right),CO_2} = \sum_{i=1}^{n} \frac{G_i(P_{i,max})}{J_{i,CO_2}\left(P_{i,max}\right)} \tag{17}$$

$$K_{i\left(\frac{max}{max}\right),CO_2} = \sum_{i=1}^{n} \frac{\omega_{0,i} + \omega_{1,i} P_{i,max} + \omega_{2,i} P_{i,max}^2 + \omega_{3,i} P_{i,max}^3}{\alpha_{0(CO_2),i} + \alpha_{1(CO_2),i} P_{i,max} + \alpha_{2(CO_2),i} P_{i,max}^2 + \alpha_{3(CO_2),i} P_{i,max}^3} \tag{18}$$

The max/max PPF (calculated in \$/kg) of the *i*th generating unit for SO_2, NO_x and CO_2 pollutant can be calculated using the above equations.

2.9 Multi-objective function

The output obtained from an optimization algorithm consists of both dominated and non-dominated solutions. Pareto dominance theory is employed to carry out this classification. For various combinations of weights, the non-dominated solution is obtained and stored in the archiving matrix. Fuzzy decision-making technique is utilized to obtain the best compromised solution. Each objective function can be converted into membership function using (19)

$$\gamma_{i,j} = \begin{cases} 1, & fn_j(P_i) \le fn_j^{min}; \\ \frac{fn_j^{max} - fn_j(P_i)}{f_{nj}^{max} - fn_j^{min}}, & fn_j^{min} \le fn_j(P_i) \le fn_j^{max}; \\ 0, & fn_j(P_i) \ge fn_j^{max} \end{cases} \tag{19}$$

where $fn_j(P_i)$ is the *j*th objective function for the *i*th solution P_i. Here, fn_j^{max} and fn_j^{min} are maximum and minimum values of the *i*th objective, respectively.

For each Pareto optimal solution *k*, the normalized membership value γ_i is calculated as explained in (20).

$$\gamma_i = \frac{\sum_{j=1}^{k} \gamma_{i,j}}{\sum_{i=1}^{nd} \sum_{j=1}^{k} \gamma_{i,j}} \tag{20}$$

where nd is the number of non-dominated solution. The function γ_i can be treated as a membership function for non dominated solution in a fuzzy set. The solution that attains the maximum membership γ_i in the fuzzy set so obtained can be chosen as the best solution (Kothari et al. 2009).

The objective functions considered in this work are as follows:

a) Fuel cost and SO_2 emission minimization.
b) Fuel cost and NO_x emission minimization.
c) Fuel cost and CO_2 emission minimization.

3 PUBG: a brief note

PUBG is one of the most widely played online multiplayer game. Like all battle royale games, PUBG also starts with some initial number of players. As the game advances, the area in the game starts to shrink. The player must protect himself from opponents as well as the shrinking zone. The last player remaining in the game is the winner. Depending upon the skill set of a player, ranks are assigned. There are total of 8 ranks in PUBG. The ranking system of the game is divided into following ranks, listed from worst to best.

Bronze → Silver → Gold → Platinum
→ Diamond → Crown → Ace → Conqueror

3.1 Methodology

In PUBG, there are number of teams. Each team comprises of number of players. Winning game needs effort from each player. Every player of the team tries to perform well to make their team win. The performance of the player depends upon his ranking. A team's performance is judged by the number of kills it has obtained. The best team in the game is the one with the most kills (highest or lowest value of objective function). The following assumptions are made to implement the proposed algorithm.

a) There are six categories of teams. The team with the highest number of kills (highest value for objective function) is considered as conqueror.
b) All other teams are chosen at random.
c) The team will have all similar ranked players, as PUBG's ranking algorithm groups all similarly rated players onto the same squad.
d) Ranks like Bronze, Silver and Gold are clubbed together into a single rank, termed as Rookie, as players in these ranks are considered to have the same skill set.
e) Shrinking of the area is also not considered as it will reduce the search space.
f) Though the game starts with 100 players (25 teams, 4 player each), in PRBO algorithm, there is no constraint in selecting the number of players or number of teams.
g) The number of players in a team can be 2 or 4, but in the proposed PRBO algorithm, number of players in a team is analogous to number of variables of the optimization problem, hence, there is no limit to number of players in a team.

Every team tries to win by improving their performance constantly. Their performance improvement equations based on the defined categories are as follow:

(i) **Conqueror:** The team with the highest overall score is considered a conqueror team. The performance of this team can be expressed as in (21)

$$X_{conquerornew,p} = X_{conqueror,p} + (\cos(\pi))^m \times rand \times \left(X_{conqueror,p} - X_{i,p}\right) \tag{21}$$

where $X_{conqueror,p}$ refers to number of kills by the pth player of the conqueror team, $X_{conquerornew,p}$ is the updated performance/number of kills by the pth player of the conqueror team, $X_{i,p}$ is the position/number of kills of randomly selected pth player of ith team in the game, *rand* is the random number that lies between [0, 1] and m is randomly taken as 1 or 2.

(ii) **Ace:** All players of this team are ace ranked players. The selection of this category is a random process. The performance of this team can be given as in (22)

$$X_{ace,p} = X_{conqueror,p} + rand \times \left(X_{conqueror,p} - X_{i,p}\right) \tag{22}$$

where $X_{conqueror,p}$ refers to number of kills by the pth player of the conqueror (best) team, $X_{ace,p}$ is the updated position/number of kills by the pth player of ace ranked team.

(iii) **Crown:** This team comprises of crown ranked player. These categories of players are slightly better than the average player of the game. The performance of this team can be expressed as in (23)

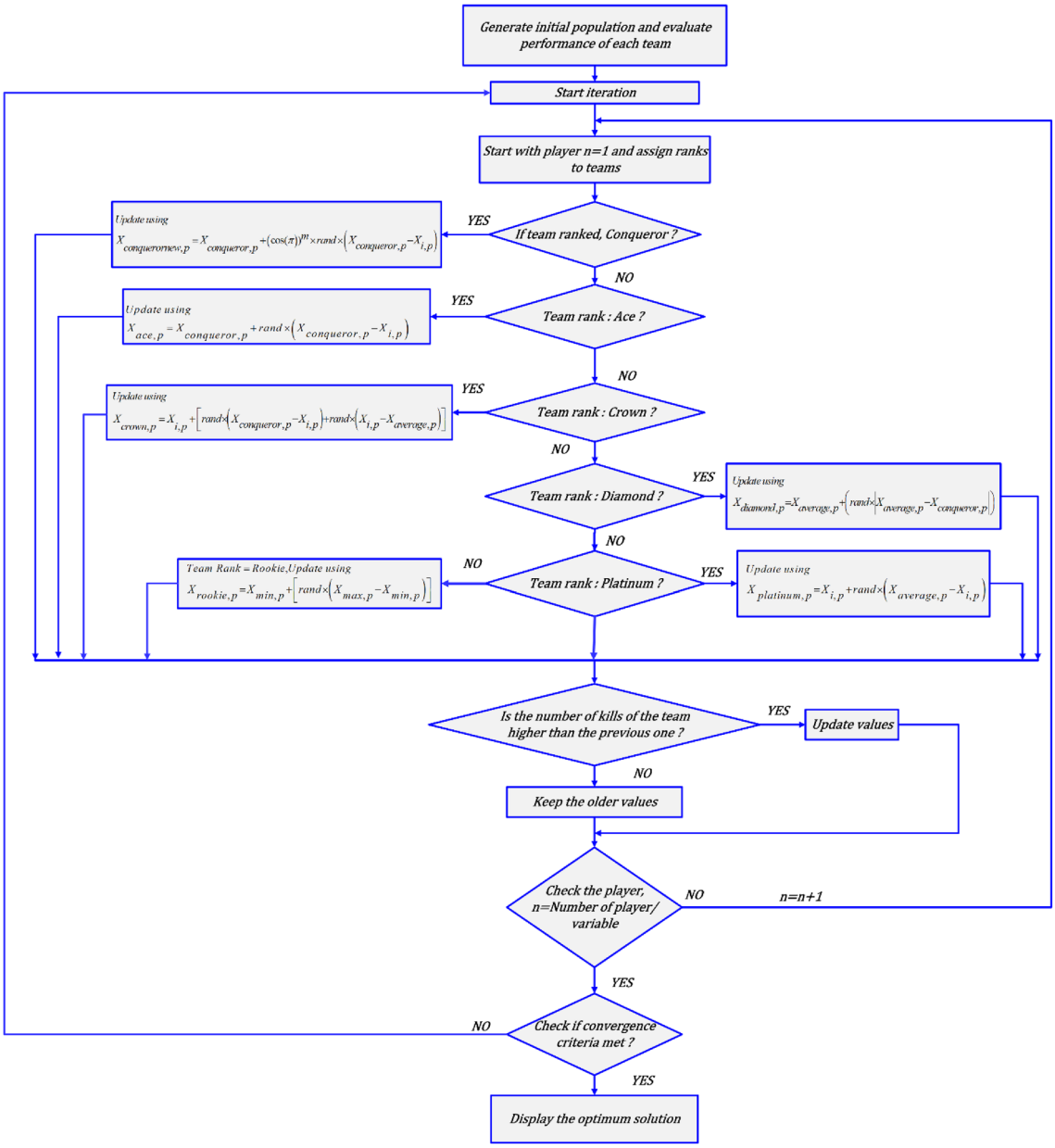


Fig. 1 The flowchart of the PRBO algorithm

$$X_{crown,p} = X_{i,p} + \left[rand \times \left(X_{conqueror,p} - X_{i,p}\right) + rand \times \left(X_{i,p} - X_{average,p}\right)\right] \tag{23}$$

where $X_{crown,p}$ is the updated position/number of kills by the pth player of the crown ranked team and $X_{average,p}$ is the average number of kills by the pth player of all the teams.

(iv) **Diamond:** This team consists of players who are diamond ranked. The performance of this team can be expressed as in (24)

$$X_{diamond,p} = X_{average,p} + \left(rand \times \left| X_{average,p} - X_{conqueror,p} \right| \right) \tag{24}$$

where $X_{diamond,p}$ is the updated position/number of kills by the ith player of the diamond ranked team and $X_{iaverage}$ is the average number of kills by the pth player of all the teams.

(v) **Platinum:** These set of players does not have the necessary expertise to win the game. These players are pushed to this category from lower ranks due to their consistent playing of game and few kills they have earned. Performance of this team can be expressed as in (25)

$$X_{platinum,p} = X_{i,p} + rand \times (X_{average,p} - X_{i,p}) \tag{25}$$

where $X_{platinum,p}$ is the updated position/number of kills by the pth player of the platinum ranked team.

(vi) **Rookie:** This rank comprises of gold, silver, and bronze category of players. These are generally new players and bots. The performance of this team can be expressed in (26)

$$X_{rookie,p} = X_{min,p} + [rand \times \left(X_{max,p} - X_{min,p} \right)] \tag{26}$$

where $X_{rookie,p}$ gives the position/number of kills update by the pth player of rookie team, $X_{min,p}$ and $X_{max,p}$ is the minimum and maximum number of kills of the pth player.

PRBO, like other population-based algorithms, starts with a population that is analogous to the number of team. Each team has a number of players that corresponds to the number of variables. To ensure that their team wins, each player must perform well. The number of kills by a team may be said as the fitness of the population. If the team's overall performance improves, the player's efforts will be acknowledged. Changes in variables will also be approved if the fitness of the population improves.

It has been considered in this algorithm that the performance of the player depends upon its designated rank. The team with the highest score is considered conqueror. Choosing of other ranks (such as ace, crown, diamond, platinum, and rookie) is a completely random process. It is noteworthy to mention that the contribution of the best team (conqueror rated) in reaching the optimal solution is significantly more than that of the other lower ranked team. However, it should be pertaining to note that all the phases contribute to achieving an optimal solution and reaching the convergence point. As a result, it can be inferred that each step plays an essential part in the algorithm's performance. All six steps are necessary to construct the PRBO, which assists the algorithm in seeking the optimal solution with the fastest convergence mobility. If any of the stages fails or is (are) eliminated, the algorithm may fail to provide the expected optimal solution or may require more number of iterations to converge. As a result, the PRBO algorithm works best when all six phases are carried out at the same time. In addition, the best solution is obtained by the conqueror ranked team, all other equations viz., ace, crown and diamond ranked team progress through the best solution obtained by the conqueror, thereby, providing necessary exploitation to the algorithm.

To guarantee that the solutions are within the feasible region, the modified population must be evaluated after each stage of modification to determine whether the populations are within the viable region or not. If any of the new population is found to be outside of the region, it must be confined within the limits. The team's performance is evaluated at every level. The team with maximum number of kills is the best team or conqueror rated team. The flowchart of the PRBO technique is shown in Fig. 1 while its pseudo code is presented in Algorithm 2.

Algorithm 2 The pseudo code of the PRBO algorithm

Step 1 Initialize the Population & Calculate the performance of each team

Step 2 **while** t<= Maximum number of iterations

for k=1 to Number of players/variables

for each team= check category

if team the best team

$X_{conquerornew,p} = X_{conqueror,p} + (cos(\pi))^m \times rand \times (X_{conqueror,p} - X_{i,p})$

else

if team category = ace

$X_{ace,p} = X_{conqueror,p} + rand \times (X_{conqueror,p} - X_{i,p})$

else

if team category = Crown

$X_{crown,p} = X_{i,p} + [rand \times (X_{conqueror,p} - X_{i,p}) + rand \times (X_{i,p} - X_{average,p})]$

else

if team category = Diamond

$X_{diamond,p} = X_{average,p} + (rand \times |X_{average,p} - X_{conqueror,p}|)$

end if

end if

end if

else

if team category = Platinum

$X_{platinum,p} = X_{i,p} + rand \times (X_{average,p} - X_{i,p})$

else

$X_{rookie,p} = X_{min,p} + [rand \times (X_{max,p} - X_{min,p})]$

end if

end if

end for

if performance better than previous performance, update values

else

Keep the previous best performance

end if

k=k+1

end for

t=t+1

end while

Step 3 Display the optimum results

4 Results and analysis

The PRBO is tested over various benchmark test functions and power system optimization problems.

4.1 Application on benchmark functions

Before applying the proposed PRBO for the CEED problem, it has been tested on 6 benchmark functions (see Table 1 for description). The research was carried out further by using the PRBO to tackle multimodal type CEC 2015 benchmark problems. Table 2 lists the considered CEC 2015 benchmark problems. The search limit for all test functions is $[100, 100]^D$. The performance of PRBPO is compared with various state-of-the-art literature as described in (Hayyolalam and Pourhaji Kazem 2020; Nasir Ghafil et al. 2022). The values of the different parameters for the chosen algorithms are tabled in Table 3.

The comparative results for the mathematical benchmark functions are shown in Table 4. The number of iterations considered for PRBO is 1000. Three population sizes are considered for the PRBO algorithm to evaluate

Table 1 Description of benchmark functions

Function number	Function name	Formula	Dimension	Range	Type
F-1	Ackley	$f_{min} = -20exp\left(-0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} x_i^2}\right) - exp\left(\frac{1}{D}\sum_{i=1}^{D} Cos2\pi x_i\right) + 20 + e$	10	[−35,35]	Multimodal non-separable
F-2	Gricwank	$f_{min} = \frac{1}{4000}\left[\sum_{i=1}^{D} x_i^2\right] - \left[\prod_{i=1}^{D} cos\left(\frac{x_i}{\sqrt{i}}\right)\right] + 1$	10	[−100, 100]	Multimodal nonseparable
F-3	Rastrigin	$f_{min} = \sum_{i=1}^{D} \left[x_i^2 - 10\cos\left(2\pi x_i\right) + 10\right]$	10	[−5.12, 5.12]	Multimodal separable
F-4	Sphere	$f_{min} = \sum_{i=1}^{D} x_i^2$	10	[−5.12,5.12]	Unimodal separable
F-5	Sum square	$f_{min} = \sum_{i=1}^{D} ix_i^2$	10	[−10,10]	Unimodal
F-6	Schwefel	$\sum_{i=1}^{D} 418.9829 - x_i sin\left(\left\vert\sqrt{\vert x_i\vert}\right\vert\right)$	10	[−500,500]	Multimodal

Table 2 CEC 2015 benchmark functions

Function number	Functions	Related basic functions	Dimension	$f_{\min}$	Type
CEC-1	Shifted and Rotated Weierstrass Function	Weierstrass Function	30	300	Simple Multimodal
CEC-2	Shifted and Rotated Schwefel's Function	Schwefel's Function	30	400	Simple Multimodal
CEC-3	Shifted and Rotated Katsuura Function	Katsuura Function	30	500	Simple Multimodal
CEC-4	Shifted and Rotated HappyCat Function	HappyCat Function	30	600	Simple Multimodal
CEC-5	Shifted and Rotated HGBat Function	HGBat Function	30	700	Simple Multimodal

Table 3 Parameters value

Algorithm	Parameter	Values
GA	Crossover rate (p_c)	0.67
	Mutation rate	0.33
	Number of iteration and population size	2000, 200
PSO	Inertia weight	2
	Best personal experience	2.4
	Best global experience	2.2
	w-damp	0.98
	Number of iteration and population size	2000, 200
BWO	procreate rate ($_{Pp}$)	0.6
	cannibalism rate (CR)	0.44
	mutation rate (P_m)	0.4
	Number of iteration and population size	2000, 200
ABC	The number of food sources	$Popsize$/2
	Limit	15
	Number of iteration and population size	2000, 200
BBO	Rate of keeping habitat	0.6
	Mutation rate	0.4
	Number of iteration and population size	2000, 200
PRBO	Population size	20, 35, 50
	Number of iteration	1000

its performance. The rank summary, presented in Table 4, shows that out of 6 benchmark functions considered, PRBO ranks first in five and third in function F-2. The simulation time for each population is also provided and it can be observed that increasing the population size does not improve the results but increases the simulation time.

The statistical results for the CEC 2015 benchmark functions have been reported in Table 5. The population size considered for PRBO is 50 in this case. The number of fitness function evaluation is considered as 1500. It can be observed that PRBO is able to obtain the lowest value for three CEC 2015 function (CEC-3, CEC-4, CEC-5) and is the second best for CEC-1 benchmark function. Additionally, PRBO complexity has been evaluated on both 10D and 30D using the guidelines specified in CEC 2015 (Chen et al. 2015). The number of fitness function evaluation is taken as 500 and 1500 for 10D and 30D function, respectively. The time required to compute the test function specified in CEC 2015 guidelines is $T_0 = 0.1342$ seconds. The algorithm complexity is given in Table 6.

The term $\frac{T_B}{T_A}$ represents the complexity of the algorithm. A value of 1 represents zero complexity from dimension (D) 10 to 30. The average value of $\frac{T_B}{T_A}$ for PRBO algorithm is 6.311 that means that the complexity for dimension 10D to 30D is increased by approximately 6 times.

Table 4 Comparative results for the considered benchmark function

Function	Parameter	GA	PSO	BWO	ABC	BBO	PRBO		
							20	35	50
F-1	B	1.15E-08	8.44E-05	5.06E-14	1.17E-01	7.27E-05	4.44E-15	4.44E-15	4.44E-15
	M	3.95–02	4.43E-03	1.57E-04	3.23E-01	1.30E-04	8.35E-15	6.93E-15	6.93E-15
	W	5.05E-01	3.15E-02	1.58E-03	9.58E-01	2.01E-04	1.51E-14	7.99E-15	7.99E-15
	SD	9.63E-02	5.95E-03	3.85E-04	1.96E-01	2.94E-05	2.62141E-15	1.71612E-15	1.71612E-15
	Rank	3	5	2	6	4	1	–	–
	Time (s)	–	–	–	–	–	3.757729	5.910856	7.512652
F-2	B	6.28E-11	5.26E-01	0	4.12E-02	7.46E-09	00	00	00
	M	4.55E-03	7.68E-01	9.43E-04	1.09E-01	2.69E-02	0.005882	.00723	0.00376
	W	2.46E-02	9.76E-01	1.26E-02	1.73E-01	9.84E-02	0.0121	0.0104	0.0084
	SD	6.85E-03	1.09E-01	2.97E-03	3.75E-02	2.36E-02	0.004132	0.00399	0.004011
	Rank	2	6	1	5	4	3	–	–
	Time (s)	–	–	–	–	–	6.045152	8.753620	12.107535
F-3	B	2.96E-11	1.47E-02	0	3.62E-08	9.95E-01	00	00	00
	M	1.07E-01	2.88E+00	4.24E-07	1.72E+00	4.74E+00	0.199	00	00
	W	6.65E-01	8.63E+00	1.94E-06	3.98E+00	8.95E+00	0.995	00	00
	SD	1.81E-01	1.82E+00	4.20E-07	1.04E+00	2.17E+00	0.419529	00	00
	Rank	3	5	2	4	6	1	–	–
	Time (s)	–	–	–	–	–	3.527586	5.536159	7.734920
F-4	B	3.3E-16	7.31E-08	4.18E-46	6.18E-05	6.95E-11	0	0	0
	M	2.58E-05	1.82E-05	2.18E-09	2.15E-04	2.27E-10	0	0	0
	W	7.27E-04	1.04E-04	5.74E-08	4.14E-04	5.04E-10	0	0	0
	SD	1.32E-04	2.84E-05	1.01E-08	9.64E-05	1.08E-10	0	0	0
	Rank	3	5	2	6	4	1	–	–
	Time (s)	–	–	–	–	–	3.171231	5.696959	5.880961
F-5	B	2.60E-12	9.59E-13	4.10E-50	9.88E-05	4.61E-09	0	0	0
	M	1.99E-03	9.04E-09	3.20E-06	7.61E-04	1.28E-08	0	0	0
	W	5.80E-02	1.62E-07	5.21E-05	1.61E-03	2.56E-08	0	0	0
	SD	1.06E-02	3.00E-08	1.04E-05	4.28E-04	4.83E-09	0	0	0
	Rank	4	3	2	6	5	1	–	–
	Time (s)	–	–	–	–	–	3.344119	5.064978	6.603853
F-6	B	4.15E+2	2.38E+02	7.50E+02	1.03E+02	4.64E+02	1.27E-04	1.27E-04	1.27E-04
	M	1.19E+3	8.73E+02	1.16E+02	3.26E+-2	9.68E+02	1.27E-04	1.27E-04	1.27E-04
	W	2.05E+3	1.43E+03	1.43E+03	4.53E+-2	1.52E+03	1.27E-04	1.27E-04	1.27E-04
	SD	3.67E+2	3.19E+02	3.19E+02	8.26E+01	2.78E+02	4.36E-09	1.07497E-11	1.04423E-11
	Rank	4	3	6	2	5	1	–	–
	Time (s)	–	–	–	–	–	2.944774	3.625538	6.226080

4.2 Application of PRBO on power system optimization problem

PRBO algorithm is applied to solve the ED problem of 3-unit, 5-unit and 26-unit systems for various loads. Additionally, the algorithm is also applied to resolve the CEED problem of IEEE 30-bus, 6-unit system for various power demands (150 MW, 175 MW, 200 MW and 225 MW). The parameters for the referred algorithm are specified in Table 7. The CCF is considered to formulate the problem. Max to max PPF is used to convert multi-objective function into single-objective function. A fuzzy set theory is adopted to determine the best compromise solution among the Pareto optimal solutions. MATLAB R2016a is utilized to simulate the PRBO to formulate the CEED problem. The simulation is carried out using Intel Celeron N4000 CPU Dual Core 1.1 GHz, with 4 GB of RAM. The algorithm is run for 30 times and the best obtained values are considered. Based on the sensitivity analysis of the algorithm, a population size of 50 is chosen for all the chosen test cases.

Table 5 Comparative results for CEC 2015 benchmark function

Function	Parameter	PSO	FF	FFPSO	FO	PRBO
CEC-1	M	3.07E+02	3.07E+02	3.14E+02	3.20E+02	3.10E+02
	SD	1.32E+00	1.24E+00	1.61E+00	7.01E-02	1.26E+00
	Rank	5	1	3	4	2
CEC-2	M	2.25E+03	1.54E+03	3.11E+03	4.86E+02	2.64E+3
	SD	3.55E+02	3.21E+02	2.62E+02	1.89E+00	1.79E+02
	Rank	3	2	5	1	4
CEC-3	M	4.74E+03	5.02E+02	5.02E+02	2.16E+03	5.00E+02
	SD	6.36E-01	5.97E-01	9.34E-01	2.03E+02	2.47E-02
	Rank	5	2	3	4	1
CEC-4	M	6.00E+02	6.00E+02	6.06E+02	1.61E+06	6.00E+02
	SD	2.84E-01	5.83E-01	1.35E+00	7.16E-10	4.03E-02
	Rank	3	2	4	5	1
CEC-5	M	7.01E+02	7.05E+02	8.05E+02	7.56E+02	7.00E+02
	SD	1.89E+00	5.80E+00	3.33E+01	6.11E+00	3.99E-02
	Rank	2	3	5	4	1

M Mean, *SD* Standard Deviation, *PSO* Particle swarm optimization, *FF* Firefly Algorithm, *FO* fertilization optimization, *FFPSO* hybrid firefly and particle swarm optimization algorithm

Table 6 Complexity observed in PRBO algorithm

Function	$D=10$		$D=30$		$\frac{T_B}{T_A}$
	T_1	$T_A=\frac{T_1}{T_0}$	T_2	$T_B=\frac{T_2}{T_0}$	
CEC-1	0.0509	0.3792	0.4525	3.3718	8.8918
CEC-2	0.0087	0.0603	0.0363	0.2704	4.4842
CEC-3	0.0270	0.2011	0.2269	1.6907	8.4075
CEC-4	0.0062	0.0461	0.0331	0.2466	5.3492
CEC-5	0.0080	0.0596	0.0354	0.2637	4.4244

ED results for 3-unit, 5-unit and 26-unit system are presented, in order, in Tables 8, 9, 10. The CEED results from PRBO for IEEE 30-bus, 6-unit system at various loads are shown in Table 11. Tables 12, 13, 14, 15, 16, 17 present the comparative analysis of PRBO with different algorithms. Table 18 displays the best compromised solution for fuel cost and different emission.

4.2.1 Case study I

The PRBO technique is implemented to solve the ED of a 3-unit and 5-unit system for 2500 MW load and 1800 MW load, respectively. The number of iterations is taken as 500. Line losses are neglected. The generation and fuel cost coefficient data are taken from (Mahdi et al. 2018).

From Table 8, it may be seen that the PRBO technique handles the ED problem efficiently. The PRBO performs better than GA and PSO technique, while its performance is marginally better than the QPSO. The convergence characteristic of PRBO for 3-unit system is illustrated in Fig. 2. It is quite evident from this figure that PRBO converges quickly, and the results are obtained within 20 iterations. The computational time for the simulation is 5.942740 s.

ED results for a 5-unit system for a load of 1800 MW is presented in Table 9. It is quite evident that PRBO technique fairs well against all the optimization techniques considered. Figure 3 shows the convergence characteristic of PRBO for this system. It can be concluded that the convergence characteristics is smooth and stable. It converges within 30 iterations. The computational time for the simulation is 6.653854 s. The number of iterations taken by PSO and GA are not given in the referred literature. QPSO technique converges within 20 iterations for both the cases. The MRFO-GBO converges in 80 iterations.

4.2.2 Case study II

In this test case, PRBO is implemented to solve the ED problem of 26-unit system to examine the viability of PRBO on a larger system. The fuel cost coefficient and power limits are taken from (Theerthamalai and Maheswarapu 2010). The number of iterations is taken as 500. The PRBO technique is compared against λ-logic, simplified direct search method (SDSM) and QPSO. Loads considered are 2000 MW and 2200 MW. The transmission line losses are not included in this case.

Table 10 presents the ED result for 26-unit system with 2000 MW and 2200 MW load. For both the cases, PRBO is optimizing the problem but lags behind other referred techniques. It should also be noted that the simulation time increases drastically with the increase in the size of power systems. The simulation time taken by PRBO for 2000 MW load is 46.645873 s while for 2200 MW load, it is 47.504226 s.

Table 7 Parameter values of different algorithms for power system optimization problem

Algorithm	Parameter	Value
QPSO	Population size	1000
	Number of runs	100
	Number of dimension	6
	Adaptive contraction factor	Linearly decreasing from 1 to 0.5
MRFA	Population size, N	30
	Number of dimension	50
	Number of runs	30
SA	Initial temperature	10 °C
	alpha	0.99
	Max tries	10,000
	Final temperature	1E-1 °C
SDSM	λ (For 2000 MW load)	13.90887
	λ (For 2200 MW load)	11.89518
LR	λ	8
PSO	Population size	100
	Inertia weight factor (w)	0.7
	Constriction factor (C_1, C_2)	− 0.2, − 0.2
QITFA	Number of dimension	6
	Gravitational constant, G	0.667
	Mass of the particles (m_1, $m_{2)}$	0.1 and 0.1
	g	Linearly decreasing from 1 to 0.5
QBA	The maximal and minimal pulse rate	1 and 0, respectively
	The maximal and minimal frequency (n_{max} and n_{min})	1.5 and 0, respectively
	The maximal and minimal loudness	2 and 1, respectively
	Delta (δ)	0.9
	Gamma (γ)	0.9
	The frequency of updating the loudness and emission pulse rate (G)	10
	The maximum and minimum probability of habitat selection	0.9 and 0.6, respectively
	The maximum and minimum compensation rate for Doppler effect (C_{max} and C_{min})	0.9 and 0.1, respectively
	The maximum and minimum contraction expansion coefficient (b_{max} and b_{min})	1 and 0.5, respectively
	The maximum and minimum inertia weight (w_{max} and w_{min})	0.9 and 0.5, respectively

Table 8 ED results for 3-unit at 2500 MW system load

Generator number	GA (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	QPSO (Mahdi et al. 2018)	PRBO [Proposed]
1	725.02	724.99	726.866	724.8267
2	910.19	910.15	908.4451	910.4761
3	864.88	864.85	864.6898	864.6972
Total power (MW)	2500	2500	2500	2500
Fuel cost ($/h)	22,730.14	22,729.35	22,728.52	22,728.41

Figures 4, 5 show the convergence characteristics of PRBO for 2000 MW and 2200 MW system, respectively. For a 2000 MW system, the PRBO converges within 350 iterations, while for 2200 MW load, it converges within 150 iterations. The λ-logic and SDSM techniques are non-iterative methods, the number of iterations taken by QPSO is approximately 20 for both the cases.

4.2.3 Case study III

The PRBO is implemented to solve the CEED of an IEEE 30-bus, 6-unit system with various loads. Max/max PPF is used to transform the multi-objective function to single-objective function. Emission of CO_2, NO_x and SO_2 are considered. The fuel cost coefficient, generator power limits,

Table 9 ED results for 5-unit at 1800 MW system load

Generator number	GA (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	QPSO (Mahdi et al. 2018)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
1	320	320.00	320.51	320	320
2	343.74	343.70	346.47	343.8698	343.8543
3	472.60	472.60	482.95	472.2678	472.2787
4	320.00	320.00	320	320	320
5	343.74	343.70	330.08	343.862	343.8669
Total power (MW)	1800	1800	1800	1800	1800
Fuel cost ($/h)	18,611.07	18,610.4	18,610.03	18,609.6904	18,609.69

Table 10 ED results for 26-unit system at different load

Generator number	Power demand = 2000 MW				Power demand = 2200 MW			
	λ-logic (Mahdi et al. 2018)	SDSM (Mahdi et al. 2018)	QPSO (Mahdi et al. 2018)	PRBO [Proposed]	λ-logic (Mahdi et al. 2018)	SDSM (Mahdi et al. 2018)	QPSO (Mahdi et al. 2018)	PRBO [Proposed]
1	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
2	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
3	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
4	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
5	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4
6	4	4	4	4	4	4	4	4
7	4	4	4	4	4	4	4	4
8	4	4	4	4	4	4	4	4
9	4	4	4	4	4	4	4	4
10	15.2	15.2	15.2	15.2	33.20	35.2	33.21	15.2
11	15.2	15.2	15.2	15.2	31.01	33.2	21.78	48.88
12	15.2	15.2	15.2	15.2	29.03	28.2	34.6	15.2
13	15.2	15.2	15.21	15.2	26.91	26.2	30.72	40.86
14	25	25	25	25	25	25	25	25
15	25	25	25	25	25	25	25	25
16	25	25	25	25	25	25	25	25
17	129.71	123.60	123.45	119.7053	155	154.25	154.98	155
18	124.70	121.25	128.06	109.7854	155	154.25	154.96	155
19	124.41	118.25	111.52	155	155	154.25	154.97	155
20	116.71	116.25	116.62	94.84937	155	154.25	154.97	155
21	68.95	68.95	68.95	68.95	68.95	68.95	68.95	68.95
22	68.95	68.95	68.95	68.95	68.95	68.95	68.95	68.95
23	68.95	68.95	68.95	68.95	68.95	68.95	68.95	68.95
24	337.69	350	349.69	350	350	350	350	350
25	400	400	400	400	400	400	399.99	400
26	400	400	400	400	400	400	399.97	400
Fuel cost ($/h)	27,246.75	26,246.42	27,245.05	27,252.37	29,725.48	29,725.32	29,724.64	29,730.89

emission coefficient and their PF are given in (Bodha, Yadav and Mukherjee, 2020). Number of iterations is taken as 500 for all the cases. The results are collated with LR, PSO, SA, QBA, QPSO, QITFA, SCA and MRFO-GBO techniques.

Table 11 displays results of CEED for various loads of 6-unit system obtained by PRBO technique. The table shows fuel cost ($/h), emission of pollutants CO_2, SO_2 and NO_x (kg/h), emission penalty factor for all the pollutants ($/h)

Table 11 CEED results using PRBO for various loads

		150.MW	175.MW	200.MW	225.MW
Power generation.(MW)	P_1	50	50	50	50
	P_2	20	24.0179	31.2531	37.7156
	P_3	15	15.0000	15.0000	17.7337
	P_4	25.56	32.8219	39.0211	44.7490
	P_5	17.55	24.0829	29.6375	34.8030
	P_6	21.90	29.0795	35.0901	40.0000
Fuel cost.($/h)		2705.6055	3190.0000	3729.2439	4315.0484
Emission.(kg/h)	CO_2	2560.59	3122.9295	3709.2186	4322.7318
	SO_2	3095.66	3804.2167	4537.4649	5288.7634
	NO_x	2405.56	2852.0632	3323.3471	3819.9342
Emission penalty cost* ($/h)	CO_2	2242.36	2731.6702	3263.4535	3848.2452
	SO_2	2726.61	3206.1746	3729.3151	4312.5686
	NO_x	2563.43	3102.4050	3685.9026	4306.8158
Total cost ($/h)		10,237.75	12,230.3048	14,407.9471	16,782.6938

*Emission penalty charge ($/h) = (Penalty factor) ×(Emission)

Table 12 Comparison of fuel cost ($/h) for 6-unit system at various loads

Load demand (MW)	LR (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	SA (Mahdi et al. 2018)	QBA (Mahdi et al. 2019)	QPSO (Mahdi et al. 2018)	QITFA (Bodha et al. 2020)	SCA (Gonidakis and Vlachos 2019)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
150	2729.35	2734.20	2705.21	2704.97	2704.89	2711.71	2704.923	2706.58	2705.60
175	3475.41	3236.30	3220.51	3188.10	3187.93	3189.91	3188.151	3206.15	3190.00
200	4210.30	3784.90	3735.73	3726.08	3727.61	3729.26	3727.446	3742.84	3729.24
225	5130.53	4402.30	4321.52	4314.63	4315.03	4316.34	4315.048	4319.33	4315.04

Table 13 Comparison of CO_2 emission (kg/h) for 6-unit system at various loads

Load demand (MW)	LR (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	SA (Mahdi et al. 2018)	QBA (Mahdi et al. 2019)	QPSO (Mahdi et al. 2018)	QITFA (Bodha et al. 2020)	SCA (Gonidakis and Vlachos 2019)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
150	2537.12	2607.10	2568.95	2565.14	2564.82	2561.85	2564.567	2562.82	2560.59
175	3613.53	3178.00	3094.69	3129.78	3129.91	3133.60	3129.187	3097.60	3122.93
200	4473.37	3771.50	3714.33	3719.64	3715.45	3721.30	3715.637	3681.42	3709.21
225	5502.52	4403.00	4324.30	4323.47	4322.77	4322.39	4322.521	4315.41	4322.73

Table 14 Comparison of SO_2 emission (kg/h) for 6-unit system at various loads

Load demand (MW)	LR (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	SA (Mahdi et al. 2018)	QBA (Mahdi et al. 2019)	QPSO (Mahdi et al. 2018)	QITFA (Bodha et al. 2020)	SCA (Gonidakis and Vlachos 2019)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
150	3091.648	3193.6	3138.446	3147.38	3146.86	3087.92	3146.831	3145.54	3095.66
175	4142.176	3904.9	3763.478	3858.96	3859.89	3807.49	3859.359	3804.23	3804.21
200	5053.584	4670.6	4553.972	4598.20	4591.38	4528.28	4592.493	4533.03	4537.46
225	6106.498	5426.1	5287.306	5344.75	5340.20	5289.45	5339.961	5318.66	5288.76

Table 15 Comparison of NO_x emission (kg/h) for 6-unit system at various loads

Load demand (MW)	LR (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	SA (Mahdi et al. 2018)	QBA (Mahdi et al. 2019)	QPSO (Mahdi et al. 2018)	QITFA (Bodha et al. 2020)	SCA (Gonidakis and Vlachos 2019)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
150	2448.218	2424.6	2379.35	2408.10	2406.37	2406.79	2406.237	2410.25	2405.56
175	2604.886	2879.7	2789.92	2853.40	2854.22	2856.22	2854.007	2838.31	2852.06
200	3102.077	3373.2	3285.647	3327.78	3324.62	3313.29	3325.339	3309.15	3323.34
225	3798.383	3877.6	3781.191	3822.53	3820.14	3815.71	3820.289	3810.87	3819.93

Table 16 Comparison of total emission (kg/h) for 6-unit system at various loads

Load demand (MW)	LR (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	SA (Mahdi et al. 2018)	QPSO (Mahdi et al. 2018)	QITFA (Bodha et al. 2020)	SCA (Gonidakis and Vlachos 2019)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
150	8076.98	8225.3	8086.742	8118.05	8056.56	8117.63	8,118.62	8061.82
175	10,360.59	9962.6	9648.086	9844.02	9797.31	9842.55	9,740.14	9779.20
200	12,629.03	11,815.3	11,553.95	11,631.45	11,562.87	11,633.47	11,523.59	11,570.03
225	15,407.40	13,706.7	13,392.80	13,483.11	13,427.55	13,482.77	13,444.95	13,431.43

Table 17 Comparison of total cost ($/h) for 6-unit system at various loads

Load demand (MW)	LR (Mahdi et al. 2018)	PSO (Mahdi et al. 2018)	SA (Mahdi et al. 2018)	QBA (Mahdi et al. 2019)	QPSO (Mahdi et al. 2018)	QITFA (Bodha et al. 2020)	SCA (Gonidakis and Vlachos 2019)	MRFO-GBO (Hassan et al. 2021b)	PRBO [Proposed]
150	10,264.57	10,385	10,261.49	10,255.28	10,255.25	10,254.91	10,255.20	10,177.05	10,237.75
175	13,251.52	12,425	12,280.04	12,241.74	12,241.71	12,241.61	12,241.66	12,134.42	12,230.30
200	16,077.41	14,642	14,421.30	14,413.88	14,413.77	14,413.56	14,413.70	14,268.09	14,407.95
225	19,661.33	17,125	16,790.69	16,783.91	16,783.86	16,783.33	16,783.78	16,599.71	16,782.69

Table 18 The best compromise solution for 6-unit system pertaining to case-1 at 225 MW system load

Unit	BCS: Fuel cost and SO_2 emission	BCS: Fuel cost and NO_x emission	BCS: Fuel cost and CO_2 emission
P_1	50	50	50
P_2	45.10	41.93	41.97
P_3	40.76	20.43	27.85
P_4	30.43	28.98	47.20
P_5	28.71	46.72	27.52
P_6	30.00	36.94	30.46
Fuel cost ($/h)	4523.587496	4421.5	4404.90
Emission (kg/h)	4709.033993	3588.4	4120.60

and the total cost ($/h) that comprise of fuel charges ($/h) and the emission cost ($/h).

Figures 6, 7, 8 and 9 show the convergence characteristics of PRBO for 150 MW, 175 MW, 200 MW and 225 MW, respectively. For 150 MW load, the PRBO converges within 30 iterations while for all other load types, it takes 100 iterations (approx.) to find the optimum solution. It is much better than LR, PSO, SA and QBA which takes 1500, 500, 10,000 and 200 iterations, respectively, to converge. QPSO and QITFA converge within 20 iterations. Number of iterations for SCA is comparable to PRBO. MRFO-GBO takes up to 300 iterations. The simulation time taken for 150 MW,

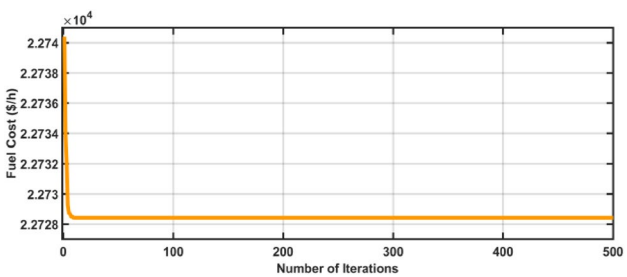


Fig. 2 The PRBO-based convergence characteristics of fuel cost for 3-unit test system

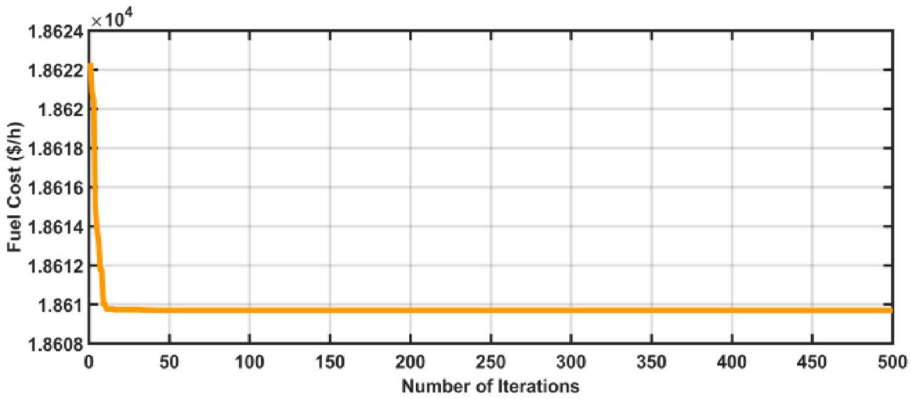


Fig. 3 The PRBO-based convergence characteristics of fuel cost for 5-unit test system

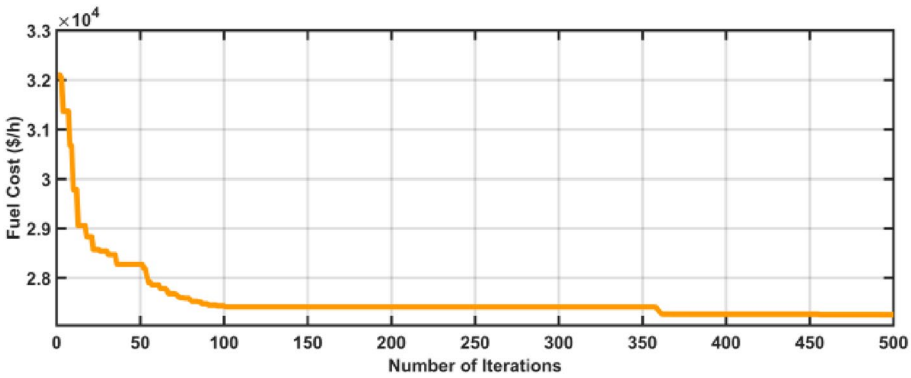


Fig. 4 The PRBO-based convergence characteristics of fuel cost for 26-unit test system corresponding to 2000 MW load

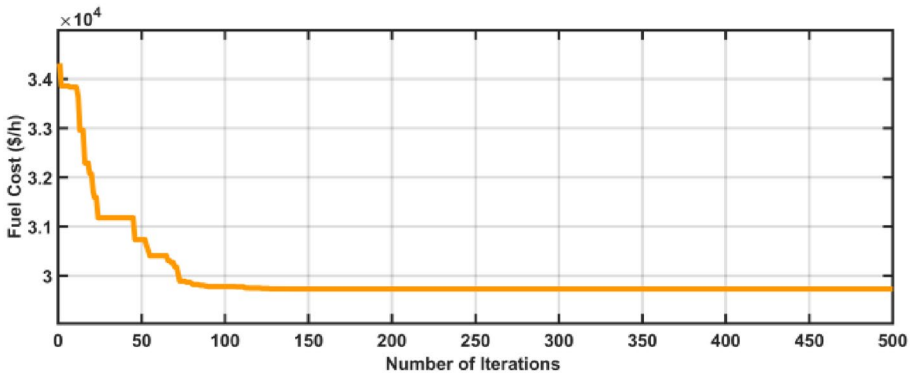


Fig. 5 The PRBO-based convergence characteristics of fuel cost for 26-unit test system pertaining to 2200 MW load

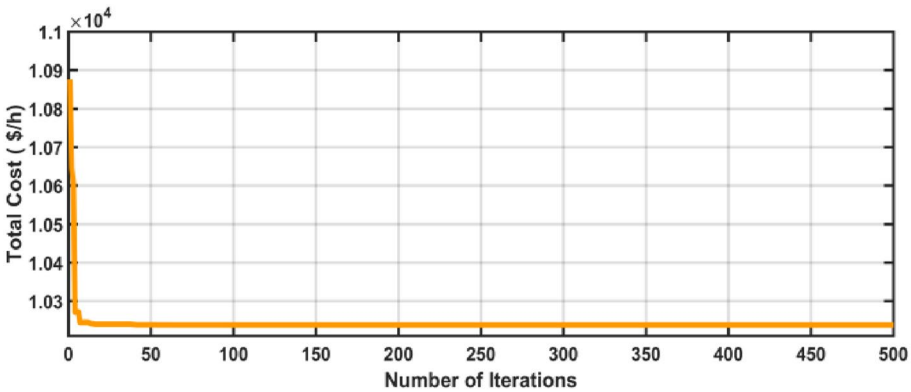


Fig. 6 The PRBO-based convergence characteristics of total cost for 6-unit test system at 150 MW load

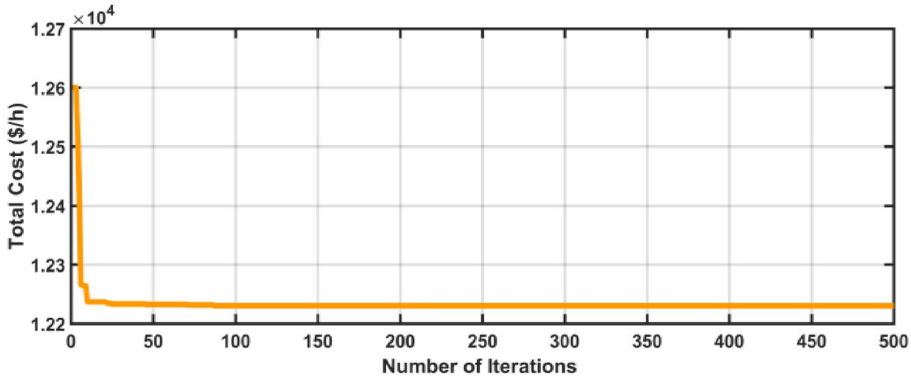


Fig. 7 The PRBO-based convergence characteristics of total cost for 6-unit test system at 175 MW load

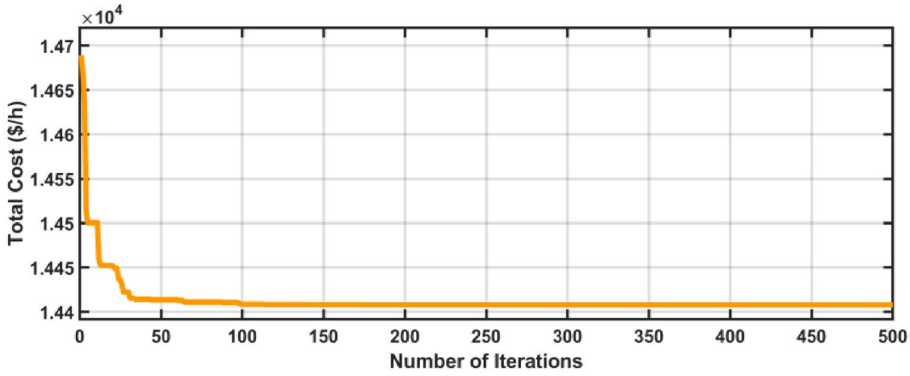


Fig. 8 The PRBO-based convergence characteristics of total cost for 6-unit test system at 200 MW load

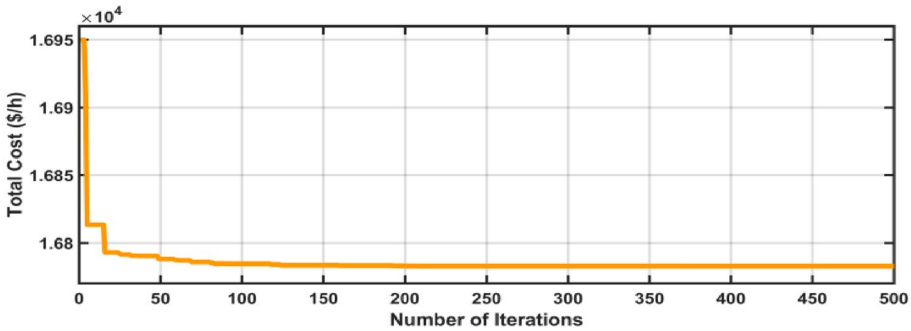


Fig. 9 The PRBO-based convergence characteristics of total cost for 6-unit test system at 225 MW load

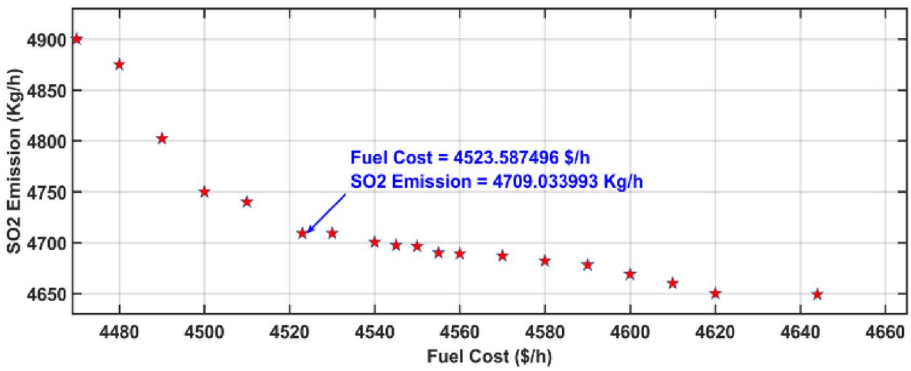


Fig. 10 The Pareto optimal front for fuel cost and SO_2

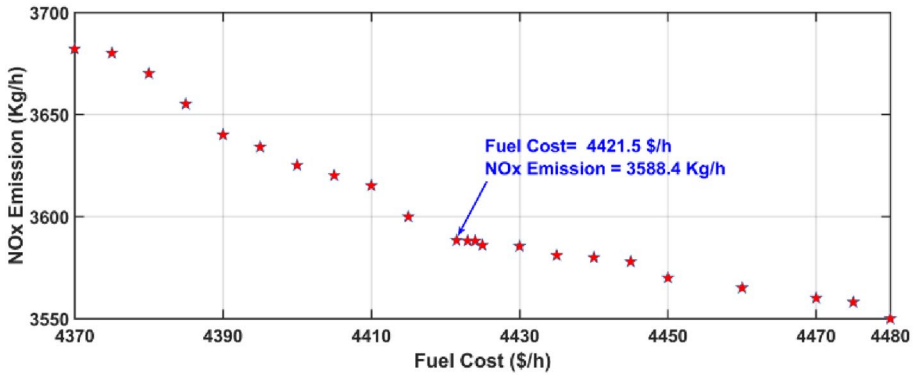


Fig. 11 The Pareto optimal front for fuel cost and NO_x

175 MW, 200 MW and 225 MW loads are 15.75 s, 16.21 s, 17.08 s, and 17.11 s respectively.

Table 12 displays the total fuel cost for different loads calculated by various optimization techniques. It can be concluded that PRBO has optimized the problem efficiently. The PRBO fares much better than classical optimization techniques like LR, SA and PSO for almost all other load types. The proposed technique provides a saving of up to 13 $/h compared to recently proposed MRFO-GBO.

Table 13 shows CO_2 emission for various techniques. For 150 MW load PRBO gives the least CO_2 emission. In case of SO_2 emission, shown in Table 14, PRBO performs better than most of the algorithms referred. Table 15 expresses the NO_x emission (kg/h) for various power demands. PRBO gives the lowest NO_x emission for 150 MW load. For other load cases, PRBO optimizes the problem well, and is among the best optimization technique taken into reference.

Total emission (kg/h) for different power demand is illustrated in Table 16. The total cost obtained by various techniques for different loads is shown in Table 17. The main objective function is to reduce the total cost and, in this case, PRBO has performed extremely well. It is evident that for all the load types, PRBO technique gives the second-best result which is just next to MRFO-GBO algorithm.

4.2.4 Case study IV

In this test case, the proposed PRBO algorithm is applied to minimize fuel cost with one the three emission types considered. This is the same system considered in case III. The optimization is done for 225 MW load.

a) **Case 1:** The PRBO algorithm is utilized to obtain best compromised values of fuel cost and SO_2 emission. Figure 10 shows the Pareto optimal values of this case. The best compromised solutions attained by the PRBO algorithm are presented in Table 18. From this table, the best compromise fuel cost ($/h) and emission of SO_2 (kg/h) are 4523.58 and 4709.03, respectively. The generating schedule is also shown in Table 18.
b) **Case 2:** In this case, best compromised solution for fuel cost and NO_x emission are considered. Figure 11 shows the Pareto optimal values for this case. The best compromised solutions obtained by PRBO algorithm are presented in Table 18. From this table, the best compromise fuel cost ($/h) and emission of NO_x (kg/h) are 4421.5 and 3588.4, respectively. Generator scheduling is also shown for this case in Table 18.
c) **Case 3:** The best compromised solution for fuel cost and CO_2 emission are considered. Figure 12 shows Pareto values for this case. The best compromised solution in this case is 4404.90 $/h and 4120.60 kg/h, as shown in Table 18.

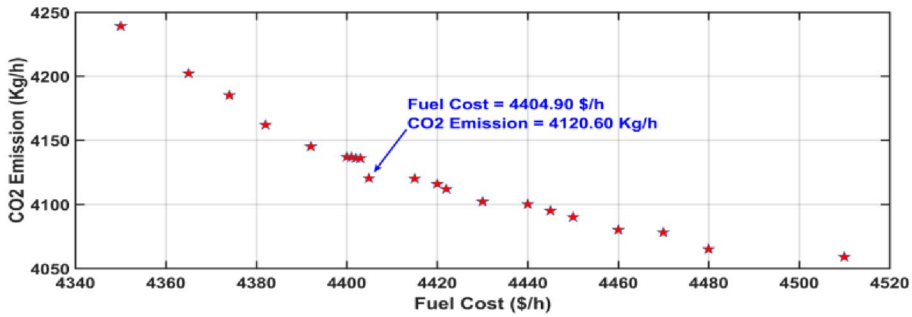


Fig. 12 The Pareto optimal front for fuel cost and CO_2

5 Conclusions and scopes of future work

The suggested PRBO algorithm is based on the popular battle royale game PUBG, in which each team strives to win by securing maximum number of kills. The performance of the team is determined by its rank. The rankings are represented by six different equations.

The PRBO algorithm is tested on 6 benchmark functions as well as 5 computationally expensive unconstrained CEC 2015 benchmark functions. The results are collated against various state-of-the-art algorithms. Additionally, the PRBO algorithm is also implemented to solve the ED problem of 3-unit, 5-unit and 26-unit system and CEED problem of IEEE 30-bus, 6-unit system. The multi-objective function is transformed into a single-objective function by utilizing unit wise max/max PPF. The obtained solutions are collated with other optimization algorithms like LR, PSO, QBA, QITFA, QPSO, SA, SCA and MRFO-GBO. The best compromised solutions for the fuel cost and various emissions are attained using fuzzy concept.

In this paper, the performance of PRBO shows that it outperforms earlier algorithms in terms of global optimal solution and convergence mobility. To solve the multi-objective function, the proposed research employs a single-objective function. The future work includes solving the CEED problem using multi-objective approach. Amalgamation of PRBO with other algorithms may be further investigated to improve its performance.

Author contributions KDB: conceptualization, Methodology, Software; VM: writing and Editing, VKY: visualization, Investigation, Supervision.

Funding The authors did not receive support from any organization for the submitted work.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Code availability The codes for the program are available at https://sites.google.com/view/prbo-optimization/home.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abdelaziz AY, Ali ES, Abd Elazim SM (2016) Flower pollination algorithm to solve combined economic and emission dispatch problems. Eng Sci Technol Int J. https://doi.org/10.1016/j.jestch.2015.11.005

Apostolopoulos T, Vlachos A (2011) Application of the firefly algorithm for solving the economic emissions load dispatch problem. Int J Combinatorics. https://doi.org/10.1155/2011/523806

Askarzadeh A, Gharibi M (2018) 'Accurate estimation of cost function parameters for thermal power plants using a novel optimization approach. Energy Sources a: Recovery, Utilization Environ Effects 40(24):2986–2999. https://doi.org/10.1080/15567036.2018.1514440 (**Taylor & Francis**)

Barbosa-Ayala OI et al (2020) Solution to the economic emission dispatch problem using numerical polynomial homotopy continuation. Energies 13(17):4281. https://doi.org/10.3390/en13174281

Benasla L, Belmadani A, Rahli M (2014) Spiral optimization algorithm for solving combined economic and emission dispatch. Int J Electr Power Energy Syst. https://doi.org/10.1016/j.ijepes.2014.04.037

Bodha KD, Yadav VK, Mukherjee V (2020) Formulation and application of quantum-inspired tidal firefly technique for multiple-objective mixed cost-effective emission dispatch. Neural Comput Appl. https://doi.org/10.1007/s00521-019-04433-0

Bodha KD, Yadav VK, Mukherjee V (2021) A novel quantum inspired hybrid metaheuristic for dispatch of power system including solar photovoltaic generation. Energy Sources B: Econ Planning Policy. https://doi.org/10.1080/15567249.2021.1933265 (**Taylor & Francis**)

Chen Q, Liu B, Zhang Q, Liang JJ, Suganthan PN, Qu BY (2015) Problem definitions and evaluation criteria for CEC 2015 special session on bound constrained single -objective computationally expensive numerical optimization. Computational Intelligence Laboratory, Zhengzhou University, China and Nanyang Technological University, Singapore, Technical Report

Chopra N, Brar YS, Dhillon JS (2021) An improved particle swarm optimization using simplex-based deterministic approach for economic-emission power dispatch problem. Electr Eng. https://doi.org/10.1007/s00202-020-01164-7

Dhillon JS, Parti SC, Kothari DP (1993) Stochastic economic emission load dispatch. Electr Power Syst Res. https://doi.org/10.1016/0378-7796(93)90011-3

Dong R, Wang S (2020) New optimization algorithm inspired by kernel tricks for the economic emission dispatch problem with valve point. IEEE Access 8:16584–16594. https://doi.org/10.1109/ACCESS.2020.2965725

El-Sehiemy RA, Rizk-Allah RM, Attia AF (2019) Assessment of hurricane versus sine-cosine optimization algorithms for economic/ecological emissions load dispatch problem. Int Trans Electr Energy Syst 29(2):1–14. https://doi.org/10.1002/etep.2716

El-Shorbagy MA, Mousa AA (2021) Constrained multiobjective equilibrium optimizer algorithm for solving combined economic emission dispatch problem. Complexity. https://doi.org/10.1155/2021/6672131

Gonidakis D, Vlachos A (2019) A new sine cosine algorithm for economic and emission dispatch problems with price penalty factors. J Inf Optim Sci 40(3):679–697. https://doi.org/10.1080/02522667.2018.1453667

Hassan MH et al (2021a) Developing chaotic artificial ecosystem-based optimization algorithm for combined economic emission dispatch. IEEE Access 9:51146–51165. https://doi.org/10.1109/ACCESS.2021.3066914

Hassan MH et al (2021b) Engineering Applications of artificial intelligence an improved manta ray foraging optimizer for cost-effective emission dispatch problems. Eng Appl Artif Intell 100(Nov 2020):104155. https://doi.org/10.1016/j.engappai.2021.104155 (**Elsevier Ltd**)

Hayyolalam V, Pourhaji Kazem AA (2020) Black widow optimization algorithm: a novel meta-heuristic approach for solving engineering optimization problems. Eng Appl Artif Intell 87:103249. https://doi.org/10.1016/j.engappai.2019.103249

He X, Rao Y, Huang J (2016) A novel algorithm for economic load dispatch of power systems. Neurocomputing. https://doi.org/10.1016/j.neucom.2015.07.107

Hossein M, Mohammad R, Siahgoli M (2020) Application of particle swarm optimization and simple intelligent algorithms in building nonlinear models for predicting formation fracture initiation pressure. Energy Sources a: Recovery Utilization Environ Effects. https://doi.org/10.1080/15567036.2020.1822466 (**Taylor & Francis**)

Jeyakumar DN, Jayabarathi T, Raghunathan T (2006) Particle swarm optimization for various types of economic dispatch problems. Int J Electr Power Energy Syst. https://doi.org/10.1016/j.ijepes.2005.09.004

Karthik N, Parvathy AK, Arul R (2019) Multi-objective economic emission dispatch using interior search algorithm. Int Trans Electr Energy Syst 29(1):1–18. https://doi.org/10.1002/etep.2683

Kothari DP, Dhillon JS (2009) Power system optimization. PHI Learning

Li Z et al (2017) Optimal state estimation for systems driven by jump-diffusion process with application to road anomaly detection. IEEE Trans Control Syst Technol 25(5):1634–1643. https://doi.org/10.1109/TCST.2016.2620062

Mahdi FP et al (2016) Emission dispatch problem with cubic function considering transmission loss using particle swarm optimization. J Telecommun Electr Comput Eng 8(12):17–21

Mahdi FP et al (2018) A quantum-inspired particle swarm optimization approach for environmental/economic power dispatch problem using cubic criterion function. Int Trans Electr Energy Syst. https://doi.org/10.1002/etep.2497

Mahdi FP et al (2019) Quantum-behaved bat algorithm for many-objective combined economic emission dispatch problem using cubic criterion function. Neural Comput Appl. https://doi.org/10.1007/s00521-018-3399-z

Mirjalili S, Hashim SZM (2010) A new hybrid PSOGSA algorithm for function optimization. In: Proceedings of ICCIA 2010 - 2010 International Conference on Computer and Information Application. https://doi.org/10.1109/ICCIA.2010.6141614

Nasir Ghafil H, Alsamia S, Jármai K (2022) Fertilization optimization algorithm on CEC2015 and large scale problems. Pollack Periodica 17(1):24–29. https://doi.org/10.1556/606.2021.00343

Pitono J (2014) Optimization of economic-emission dispatch by particle swarm optimization (PSO) using cubic criterion functions and various price penalty factors. Astra Salvensis 6(3):749–762

Precup R-E et al (2021) Optimal tuning of interval type-2 fuzzy controllers for nonlinear servo systems using Slime Mould Algorithm. Int J Syst Sci. https://doi.org/10.1080/00207721.2021.1927236

Qian S, Wu H, Xu G (2020) An improved particle swarm optimization with clone selection principle for dynamic economic emission dispatch. Soft Comput 24(20):15249–15271. https://doi.org/10.1007/s00500-020-04861-4

Radosavljević J (2016) A solution to the combined economic and emission dispatch using hybrid PSOGSA algorithm. Appl Artif Intell 30(5):445–474. https://doi.org/10.1080/08839514.2016.1185860

Rajagopalan A et al (2019) Chaotic self-adaptive interior search algorithm to solve combined economic emission dispatch problems with security constraints. Int Trans Electr Energy Syst 29(8):1–26. https://doi.org/10.1002/2050-7038.12026

Ryu H-S, Kim M-K (2020) Combined economic emission dispatch with environment-based demand response using WU-ABC algorithm. Energies 13(23):6450. https://doi.org/10.3390/en13236450

Sakthivel VP, Suman M, Sathya PD (2021) Combined economic and emission power dispatch problems through multi-objective squirrel search algorithm. Appl Soft Comput 100:106950. https://doi.org/10.1016/j.asoc.2020.106950

Shaw B, Mukherjee V, Ghoshal SP (2012) A novel opposition-based gravitational search algorithm for combined economic and emission dispatch problems of power systems. Int J Electr Power Energy Syst. https://doi.org/10.1016/j.ijepes.2011.08.012

Theerthamalai A, Maheswarapu S (2010) An effective non-iterative "λ-logic based" algorithm for economic dispatch of generators with cubic fuel cost function. Int J Electr Power Energy Syst 32(5):539–542. https://doi.org/10.1016/j.ijepes.2009.11.002 (**Elsevier Ltd**)

Xia A, Wu X, Bai Y (2021) Hybrid MHHO-DE algorithm for economic emission dispatch with valve-point effect. Arab J Sci Eng. https://doi.org/10.1007/s13369-020-05308-6

Xin-gang Z et al (2020) Economic-environmental dispatch of microgrid based on improved quantum particle swarm optimization. Energy. https://doi.org/10.1016/j.energy.2020.117014 (**Elsevier Ltd**)

Zhang J et al (2021) An improved symbiosis particle swarm optimization for solving economic load dispatch problem. J Electr Comput Eng 2021:1–11. https://doi.org/10.1155/2021/8869477 (**Edited by F Vallée**)

Ziane I, Benhamida F, Graa A (2017) Simulated annealing algorithm for combined economic and emission power dispatch using max/max price penalty factor. Neural Comput Appl. https://doi.org/10.1007/s00521-016-2335-3

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

A reformulation of collision avoidance algorithm based on artificial potential fields for fixed-wing UAVs in a dynamic environment

Astik Srivastava*. PB Sujit.**

**Delhi Technological University, New Delhi, 110042 India (Tel: 9311784289; e-mail: astiksrivastava_ep20a18_49@dtu.ac.in).*

***Indian Institute of Science, Education, and Research, Bhopal, Madhya Pradesh, India (e-mail: sujit@iiserb.ac.in)*

This work has been submitted to IFAC for possible publication

Abstract: As mini UAVs become increasingly useful in a civilian work domain, the need for a method for them to operate safely in a cluttered environment is growing, especially for fixed-wing UAVs as they are incapable of following the stop-decide-execute methodology. This paper presents preliminary research to design a reactive collision avoidance algorithm based on the improved definition of the repulsive forces used in the Artificial potential field algorithms to allow feasible and safe navigation of fixed-wing UAVs in cluttered, dynamic environments. We present simulation results of the improved definition in multiple scenarios, and we have also discussed possible future studies to improve upon these results.

Keywords: UAV, obstacle avoidance, Artificial Potential Field, Path-planning

1. INTRODUCTION

Unmanned Aerial Vehicles (UAVs) have been a vital asset for most militaries worldwide for the past two decades. In recent years, they have also started attracting significant attention in the civilian domain due to their potential for cost-effectively solving problems like organ transportation and agricultural mapping.

As these UAVs become a more integral part of the day-to-day lives of an average person, the airspace, especially in urban areas, will become cluttered and difficult to manage. Cluttered airspace also means that the probability of inter-vehicular collisions and collisions of UAVs with other objects in the environment will be high.

The Artificial Potential Field Approach (APFA) was first proposed by *Khatib et al. (1986)*. It is one of the most promising approaches for solving the collision avoidance problem due to its mathematical simplicity and computational efficiency, but the algorithm has its limitations.

The first drawback is the lack of a guarantee that the path taken by the UAV is optimum. The second and, in most cases, most catastrophic drawback is the tendency of the agent to get stuck in a local minimum, which is generally the case in environments with complex geometry or clutter.

In this paper, we propose our version of the potential field algorithm that has proven to minimise the probability of encountering a local minimum and is also inclined towards providing a more optimum path. We achieved this by reformulating the repulsive potential function that is used in prior works. The novel potential function introduced in this paper is based on the **Cauchy distribution.** The contour of the repulsive potential function in the x-y plane is of the shape of an ellipse, contrary to the circular contours seen in typical APFA (Artificial Potential Field Approaches), whose major axes always point in the direction of the relative velocity of the oncoming agent (UAV in this case) w.r.t the obstacle. We suggest a strategy to tune the gains of the potential function based on the likelihood of collision with an obstacle. The main advantage that this reformulation has demonstrated is that it generates feasible commands for a fixed-wing UAV while minimising the deviation from the desired path. To the best of our knowledge, only work in robotic literature that uses a similar concept with a different mathematical formulation is done by *TJ Stastny et al*.

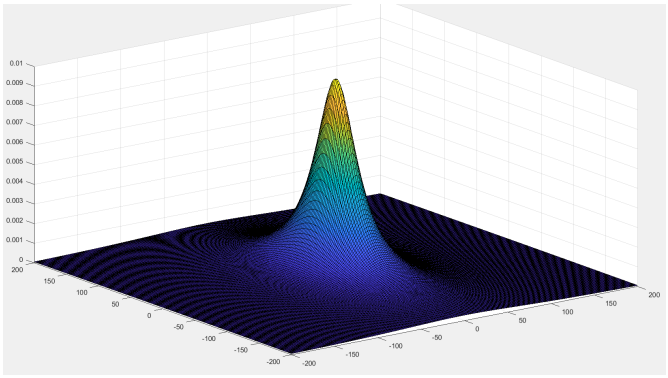
The current work is limited to solving the obstacle avoidance problem in a two-dimensional environment, although the algorithm can be extended to a 3D environment with some modifications.

2. Algorithm Formulation

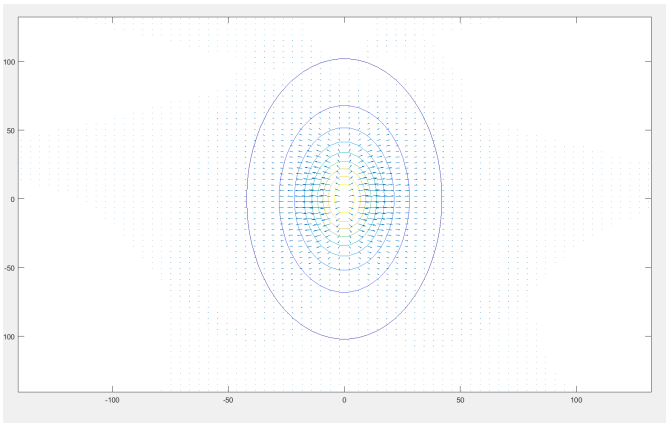
2.1 Problem Formulation

The algorithm proposed in this paper aims at performing reactive collision avoidance with minimal deviation from the path taken by the UAV. Since most mini UAVs lack control surfaces to decelerate quickly, the airspeed of the UAV is assumed to be a constant. Also, the current work aims at performing obstacle avoidance and trajectory following in a 2-dimensional environment. Hence the altitude is also kept constant.

The algorithm is primarily designed for the safe operation of fixed-wing UAVs, as generally, they move at high velocities and have limited manoeuvrability. Obstacle avoidance is achieved by constructing a potential field around an obstacle with an elliptical contour in the x-y plane, as shown (Fig. 1).



(a)



(b)

Figure 1. (a) Potential field generated for an obstacle placed at the origin, with a = 14 and b = 34. (b) gradient and contour plot of the repulsive potential function.

The major axis of the contours is parallel to the relative velocity between the obstacle and the UAV. The parallelism allows the UAV to start taking small manoeuvres when it is far from the obstacle rather than making sharp manoeuvres when it is near it. This guidance strategy reduces the perpendicular deviation of the UAV from the intended path, providing some optimality and ensuring that the manoeuvres commanded to the UAV are within its feasible range.

2.2 Guidance model of the UAV

The equations of motion for the fixed-wing UAV used are shown below(1):

$$\dot{x} = V_g * cos(\psi)$$

$$\dot{y} = V_g * sin(\psi) \qquad (1)$$

$$\dot{\psi} = \frac{g}{V} * tan(\phi)$$

This model is based on the assumption that the UAV is flying in a coordinated-turn condition, i.e., side slip angle is 0 and has been implemented using MATLAB's inbuilt Fixed-wing guidance model. More intricate control schemes like MPC, LQG controllers, etc., can be implemented to track the heading commands given by the obstacle avoidance algorithm (see *Stastny et al.(2013)*), but they are beyond the scope of this paper.

2.3 Potential field-based obstacle avoidance

Potential field algorithm works by using the gradients of a scalar potential field to guide the agent towards the desired point or along a desired path. This method is analogous to the behaviour of charges in an electric field. The agent and obstacles are analogous to like-charges, whereas the agent and the goal are analogous to opposite charges. The resulting forces then guide the UAV away from obstacles and towards the goal.

The net potential and net force experienced by the agent can be represented as shown in (2):

$$U_{net} = \sum_{i=1}^{n} U(i)_{rep} + U_{att} \qquad (2)$$

$$F_{net} = \sum_{i=1}^{n} \nabla U(i)_{rep} + \nabla U_{att}$$

The acceleration itself is not used to guide the UAV. Instead, the heading is changed to point towards the direction of the resultant force while keeping the speed constant (3).

$$\psi = tan^{-1}\left(\frac{F_y}{F_x}\right) \qquad (3)$$

2.4 Attractive potential fields

The attractive potential field to guide the UAV towards a point is defined as (4)

$$U_{att} = k_{att} * ((x - x_{wp})^2 + (y - y_{wp})^2) \qquad (4)$$

where k_{att} is a tunable parameter, (x_{wp},y_{wp}) are the coordinates of the target waypoint, and (x,y) is the position of the agent. The force due to this potential field is given by

$$F_{att} = 2 * k_{att} * [(x - x_{wp})\,\hat{i} + (y - y_{wp})\,\hat{j}] \qquad (5)$$

Although this formulation is effective in guiding the UAV towards a point, there is no fixed path that the UAV takes to reach that point when using this force. The UAV can be constrained to follow a straight/curved path from point A to point B or follow an orbit around a point with a given radius based on vector fields given by *Nelson et al. (2007).*

2.5 Repulsive potential field

The potential function around the obstacle is designed to allow smooth manoeuvring. It also minimises the deviation from the desired path while ensuring that the minimum distance between the obstacle and UAV is always greater than the *minimum-allowable separation*. A 3D Cauchy distribution with an elliptical base is used as a template for the potential function(fig 1.). The function is subject to rotation transformation about the z-axis so that the major axis of the potential field is parallel to the relative velocity vector of the incoming UAV w.r.t obstacle.(fig 2).

The potential function is given as follows (7):

$$U_{rep} = \sum_{i=0}^{n} \frac{K_{rep(i)}}{1 + (\frac{X}{a(i)})^2 + (\frac{Y}{b(i)})^2} \qquad (7)$$

where:

$$X = (x_{pos} - x_{obs(i)}) * cos(\theta_i) + (y_{pos} - y_{obs(i)}) * sin(\theta_i)$$

$$Y = (x_{pos} - x_{obs(i)}) * sin(\theta_i) - (y_{pos} - y_{obs(i)}) * cos(\theta_i)$$

$$\theta_i = tan^{-1}(\frac{V_{y(i)}}{V_{x(i)}})$$

here, $k_{rep(i)}$ is the amplitude of the function, (x_{pos}, y_{pos}) is the position of the UAV, $(x_{obs(i)}, y_{obs(i)})$ is the position of the i^{th} obstacle, and θ_i is the inclination of the relative velocity of the i^{th} obstacle w.r.t UAV, from the x-axis. In this formulation, only the length of the semi-minor axis of the ellipse is fixed and depends on the dimensions of the obstacle (or can be set to the desired minimum-safe-distance required between the UAV and the obstacle). Parameters a and k_{rep} dynamically change according to the scenario, as will be demonstrated.

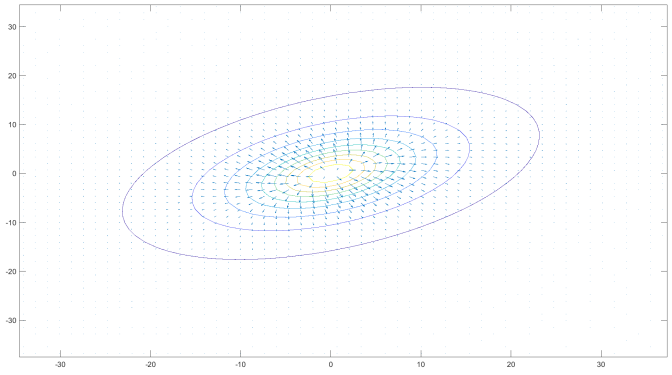


Figure 2. Potential field inclined at an angle of π/6 with a = 8 and b = 6.

We obtain the gradient of the potential field by multiplying the gradients of the non-rotated function (with major and minor axes along the x-axis and y-axis, respectively) with the transformation matrix given by

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

But, the final vectors which give the heading to the UAV are slightly different from the mathematical gradient of the potential field.

The result can be expressed as (8):

$$F_{r(i)} = F_{rep(i)}\hat{i} + F_{rep(i)}\hat{j} \qquad (8)$$

$$F_{r(i)}\hat{i} = -2 * k_{rep(i)}(u_{(i)})^{-2} * (\frac{x_{pos} - x_{ob(i)}}{a(i)^2} * cos(\theta(i)) - \frac{y_{pos} - y_{ob(i)}}{b(i)^2} * sin(\theta(i)))$$

$$F_{r(i)}\hat{j} = -2 * k_{rep(i)}(u_{(i)})^{-2} * (\frac{x_{pos} - x_{ob(i)}}{a(i)^2} * sin(\theta(i)) + \frac{y_{pos} - y_{ob(i)}}{b(i)^2} * cos(\theta(i)))$$

where î and ĵ are unit vectors along x and y-axis respectively. The coefficient $u_{(i)}$ is defined as

$$u_{(i)} = 1 + \Gamma * ((\frac{x_{pos} - x_{ob(i)}}{a(i)})^2 + (\frac{y_{pos} - y_{ob(i)}}{b(i)})^2) \qquad (9)$$

The constant Γ is introduced to shift the maxima of the magnitude of force F_{rep}. This constant is fixed to **0.34** so that the maxima lie on the minor axis, at a distance of b from the obstacle.

Parameters a_i and $k_{rep(i)}$ determine how far from the obstacle will the UAV start taking the evasive manoeuvres and how far will the UAV be deflected from its intended path, respectively. For the algorithm to produce desirable results, these two parameters vary depending on the likelihood of collision with the obstacle. This is done as follows:
We define η_i as the angle between the position vector from UAV to obstacle(i) and the relative velocity vector, then

$$\xi_i = cos(\eta_i) = \frac{(x_{obs(i)} - x_{UAV}) * cos(\theta_{(i)}) + (y_{obs(i)} - y_{UAV}) * sin(\theta_{(i)})}{\sqrt{(x_{obs(i)} - x_{UAV})^2 + (y_{obs(i)} - y_{UAV})^2}}$$

$$k_{rep(i)} = k_{rep0} * (sin(\frac{\pi * \xi_i}{2}) + 1) \qquad (10)$$

$$a_{(i)} = b_{(i)}(1.734 + \xi_i)$$

Variation of $k_{rep(i)}$ has been shown in fig.3. As you can see, when the collision is head-on, i.e. η = 0, then the amplitude of repulsive potential is maximum, and then decreases as the angle increases.

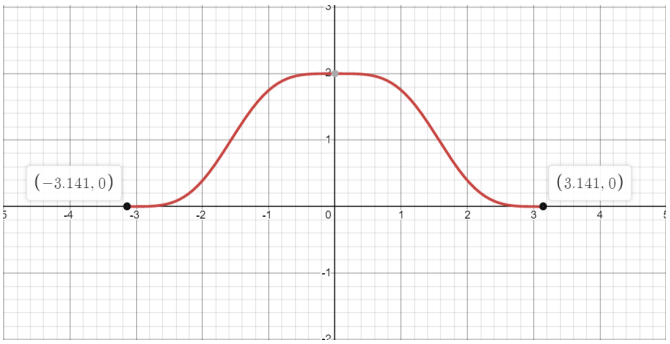


Figure 3. variation of k_{rep} (y-axis) with angle η (x-axis, between -π to π)

The coefficient k_{rep0} has been tuned heuristically for head-on collisions.

This methodology of adapting the eccentricity of contours and amplitude works well in cluttered environments. More complex adaptive algorithms based on the likelihood of collision derived using the collision cone algorithm (see *Mujumdar A (2011)*) can be used.

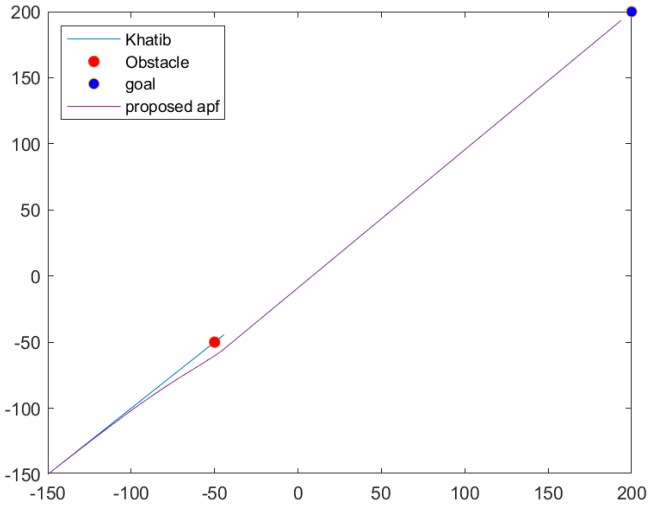
3. Simulation results and discussion

To test the algorithm's validity, multiple scenarios were constructed with both static and dynamic obstacles.

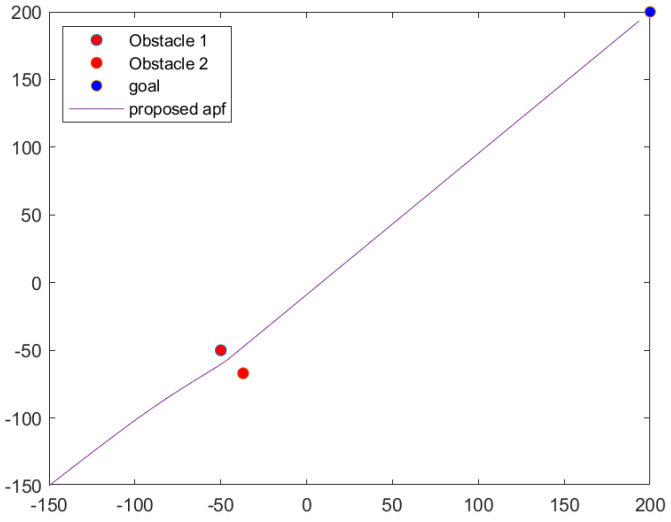
3.1 Simulation 1

The first simulation is a set of scenarios designed to test the capability of the proposed potential field algorithm in a head-on collision and when passing a narrow space between two obstacles, wherein traditional APFA tends to get stuck in local minima.

The case of Goal-Not-Reachable, where the agent is unable to reach the goal as it is close to an obstacle, is not considered as that situation is unlikely to be encountered in a real-life.



(a)

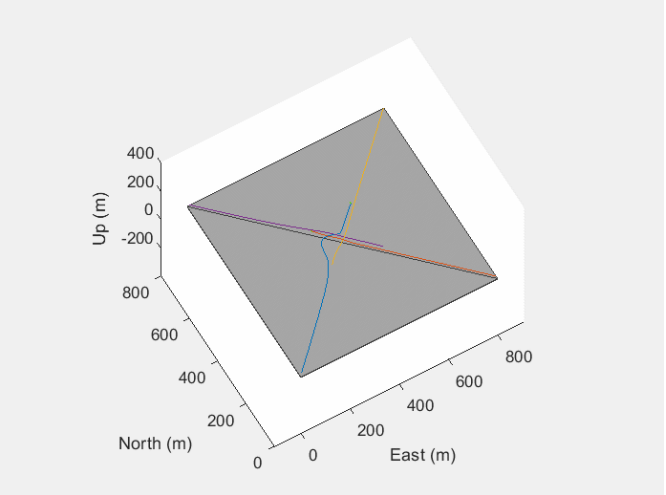


(b)

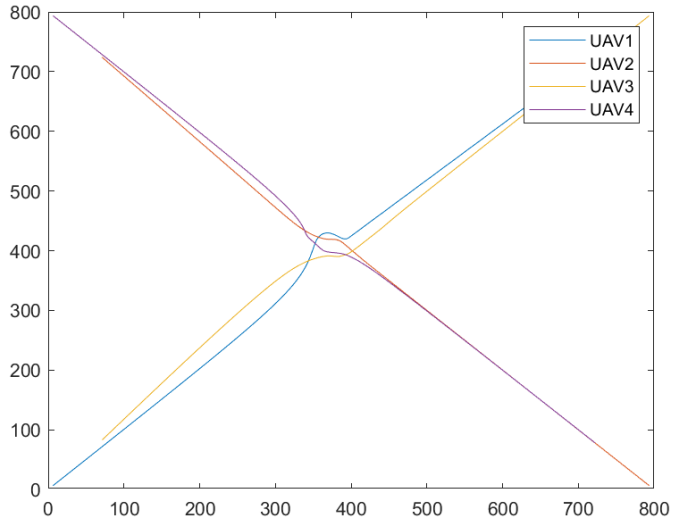
Figure 4. (a) Scenario showing the path taken by UAV with traditional APFA and with proposed APFA for a head-on collision. Path taken by the UAV in the presence of two obstacles closely spaced.

3.2 Simulation 2

The second simulation demonstrates dynamic obstacle avoidance in a multi-agent environment. Each agent independently tries to avoid the other agents and reach its goal. The goal for each UAV is the initial position of the UAV diagonally opposite, causing all agents to converge in the center. The results of this scenario are given below.



(a)



(b)

Figure 5. Path taken by 4 agents to collaboratively avoid collision

Table 1 Potential function parameters

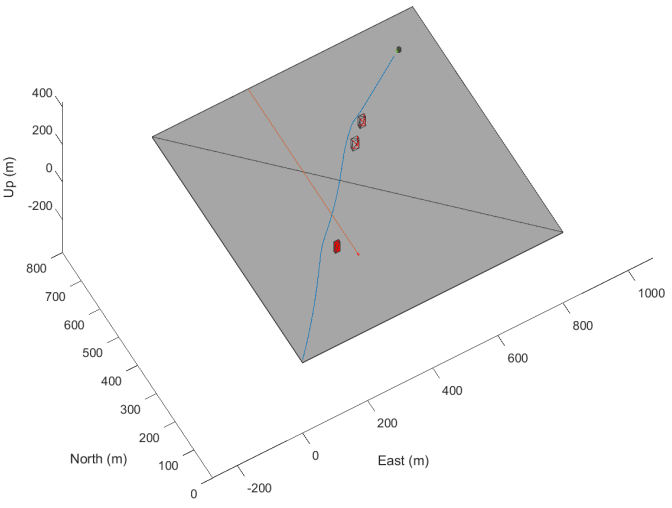
Parameters	Value	Units
Start position	UAV1 = [0, 0] UAV2 = [800,0] UAV3 = [800,800] UAV4 = [0,800]	meters
Goal position	UAV1 = [800,800] UAV2 = [0,800] UAV3 = [0,0] UAV4 = [800,0]	meters
k_{att}	0.008	-
K_{rep0}	30	-
b	[30, 30, 30, 30]	meters

3.2 Simulation 3

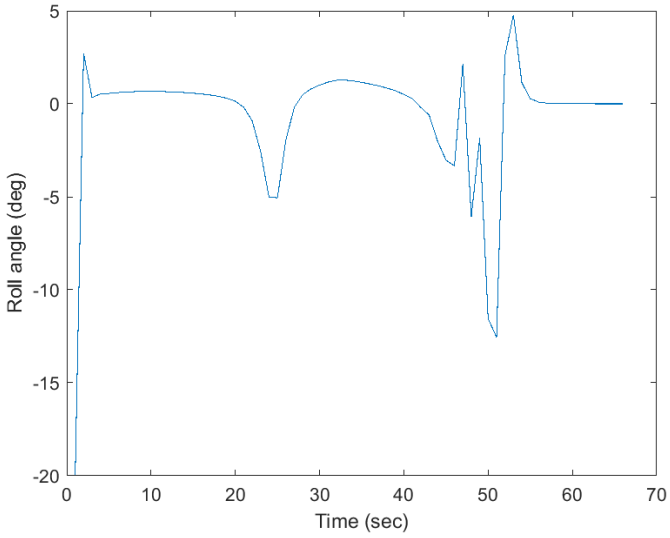
This simulation scenario has been simulated using full lateral kinematics of the UAV as shown in (1). MATLAB's UAV Toolbox has been used to design the scenario and incorporate equations (1) in the system. The UAV attempted to reach the goal while avoiding three buildings and one UAV moving with a constant velocity. The path taken by UAVs and their roll angle response are shown in figure 5. The parameters of the scenario are shown in Table 2.

Table 2 Simulation parameters

Parameters	Value	Units
Start position	(0, 0)	meters
Goal position	(800, 800)	meters
k_{att}	0.008	-
K_{rep0}	30	-
b for static obstacles	[11; 6; 9]	meters
b for dynamic obstacles	[20]	meters
Command Airspeed of UAV (V)	15	meters/sec
positions of static obstacles.	[500, 550; 450, 500; 250, 250]	meters
V_{ob2} (velocity of obstacle 1)	(0,-10)	meters/sec
Co-ordinate frame convention	East-North-UP	



(a)



(b)

Figure 6. (a) Scenario with the red path being that of obstacle UAV, and blue being that of the controlled UAV. (b) The roll angle response of the UAV.

6. CONCLUSIONS

This paper attempts to develop a new expression for the repulsive force used in the APF approach to guide UAVs. The modified algorithm was tested successfully for scenarios wherein local minima can often be encountered, like head-on collisions or passing through narrow pathways without the use of virtual forces or gyroscopic forces (see *Lee et al (2003)* and *chang et al (2003))*. This was achieved by dynamically changing the repulsive gain and the eccentricity of the elliptical contour of the potential field, but for scenarios involving a large number of moving obstacles (100-200) in a tight space, a significant jittering or oscillation was recorded. The jittering experienced in such a scenario can be improved by the dynamic step adjustment method (see *Sun et al. (2017))* and the performance will be improved in future studies.

REFERENCES

Khatib, O., 1986. Real-time obstacle avoidance for manipulators and mobile robots. In *Autonomous robot vehicles* (pp. 396-404). Springer, New York, NY.

Stastny, T.J., Garcia, G.A. and Keshmiri, S.S., 2015. Collision and obstacle avoidance in unmanned aerial systems using morphing potential field navigation and nonlinear model predictive control. *Journal of Dynamic Systems, Measurement, and Control*, *137*(1).

Chang, D.E. and Marsden, J.E., 2003. Gyroscopic forces and collision avoidance with convex obstacles. In *New trends in nonlinear dynamics and control and their applications* (pp. 145-159). Springer, Berlin, Heidelberg.

Sun, J., Tang, J. and Lao, S., 2017. Collision avoidance for cooperative UAVs with optimized artificial potential field algorithm. *IEEE Access*, *5*, pp.18382-18390.

Nelson, D.R., Barber, D.B., McLain, T.W. and Beard, R.W., 2007. Vector field path following for miniature air vehicles. *IEEE Transactions on Robotics*, *23*(3), pp.519-529.

Stastny, T.J., Garcia, G.A., Keshmiri, S.S., Ewing, M.S., Lykins, R. and Hale, R.D., 2013. Real-time collision and obstacle avoidance in unmanned aerial systems.

Mujumdar, A. and Padhi, R., 2011. Reactive collision avoidance of using nonlinear geometric and differential geometric guidance. *Journal of guidance, control, and dynamics*, *34*(1), pp.303-311.

Lee, M.C. and Park, M.G., 2003, July. Artificial potential field-based path planning for mobile robots using a virtual obstacle concept. In *Proceedings 2003 IEEE/ASME international conference on advanced intelligent mechatronics (AIM 2003)* (Vol. 2, pp. 735-740). IEEE.

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

A Review on Opinion Leader Detection and its Applications

Sonia
Department of Information and Technology
Delhi Technological University
Delhi, India
{er.soniyachoudhary}@gmail.com

Kapil Sharma
Department of Information and Technology
Delhi Technological University
Delhi, India
{kapil}@ieee.com

Monika Bajaj
Department of Computer Science
University of Delhi
Delhi, India
{mbajaj}@du.ac.in

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835870

Abstract- In today's era internet has evolved as a means for users to connect and share their opinions. So internet users can easily access information, put across their opinions, and connect with others through a wide range of social media and services. A powerful individual who is an expert in a given subject and has a significant number of followers whose opinions influence others on social media is known as an opinion leader (OL). Opinion leaders are those who are motivated, enthusiastic, and dynamic, with good educational and behavioural edifications, and who use their expert knowledge to influence clients. Opinions and ideas are the most important sensory variables in sociological events. The objective of the paper was to provide an introduction to OL detecting methods. To accomplish this, the proposed research work has discussed about the difficulty of identifying OLs followed by an evaluation of several solutions analyzed methodologies, such as centrality and community-driven metrics, methods that concentrate on interaction data and influence diffusion processes within a network, and the substance of information shared by individuals. The techniques, models, and processing flows, though, are strongly reliant on the problematic area and the unique data accessible; as a result, it is extremely challenging to create targeted interventions. In practice, however, the purpose is to detect leaders on certain issues and in definite situations, necessitating the use of the most effective specially adjusted methodologies. The methods offered in this paper reflect these considerations to choosing a suitable method or approach for a given circumstance.

Keywords- Opinion Leader, Social Network, Ant Colony Optimization, Opinion Leader Selection, Ant Generation.

I. INTRODUCTION

The majority of research on online public perception[1], comprising online opinion mining, and distribution patterns, is primarily focused on the transmission of online communal thought on social networks (SN). Users' standing in public communities is uneven; those in the middle perform a major and propelling part in the progression of online public opinions, whereas others on the perimeter are simply swayed by additional causes[2]. Online opinion leaders are frequently the core, and their engagement can quickly propel particular measures to the front Online OL are frequently at the middle, and their engagement can quickly propel particular events to the forefront of community belief. Public OL has the power to both vigorously steer social life and to elicit a variety of unpleasant feelings[3][4]. As a consequence, identifying and mining public OL is critical for directing the right public opinion and monitoring network stability.

The technique of grouping public operators into closely linked and extremely significant categories so that each category may be fine isolated from the others is known as community detection. The domains of SN research, spatial information technology, and smart grids all benefit from community detection[5][6]. This study employs communal exposure to create a communal opinion loop of individuals in social networks who share similar ideas and perspectives. The semantic cohesion[7] and structural compactness of the society are employed to additionally boost the implication of OL in the situation, rather than merely recognizing the opinion communities by leveraging links among nodes[8].

Problem Definition

These methods have several disadvantages: On the one hand, self-reported data is subjective by definition, and self-claimed relevance is likely a reflection of self-assurance.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

Having an objective evaluation of opinion leaders would result in more dependable techniques for identifying leaders. On either hand, sociological studies are restricted by the resources required to conduct them; having scalable algorithms that can properly identify OL enables us to make utilization of the ever-increasing amounts of social data enabled by the Internet. Consequently, tools based on network science and data mining have been developed to identify OL; this article focuses on these strategies.

1.1 Theoretical Basis and Concept Definition of OL

A. Network Opinion Leader

Individuals who have the potential to lead and play a key role on the Digital platform are known as network thought leaders. They are activists who frequently supply information, thoughts, remarks, and influence to individuals in interpersonal interaction channels, as well as mediators in the information chain formed by interpersonal communication[9]. In comparison to other groups, network opinion leaders are more informed, fluently skilled, politically savvy, and engaged[10]. They build social interaction linkages in a variety of ways, including self-identity, information involvement, and mutuality[11]. OL mining is an essential matter in virtual network analysis that has a lot of implications for controlling public opinion on the internet and disseminating data. To find and study opinion leaders in the system, several scholars have used information sources, k-clique grouping[12], PageRank index[13], network analysis[14], and other methods.

Members in the knowledge Q&A forum, unlike other social networks, primarily gather material or expertise from satisfied developers through research, inquiry, and conversation. As a result, this research work has focused on interrelatedness and conversation between users when recognizing online OL in the knowledge Q&A forum, such as whether the viewer frequently provides data, opinions, and remarks to other workers in the discussion community and whether the user has a consequence on other users.

B. Knowledge Q&A Community

Although there are many different types of modern popular social networks, knowledge Q&A groups have their distinct qualities. To begin, these groups are classified into two types: knowledge groups and social groups. Knowledge groups share ideas and insights through queries, whereas social groups are mostly formed through interactivity and offline salon events[15], and both of these types of groups have gained acceptance in recent years. In the United States, Quora, and Stack Overflow are samples of knowledge Q&A groups, Knowledge-IN is an example; while in China, Baidu Know, Himalaya FM, and Zhuihu are examples. The community has appreciated and is actively involved in the flow of information in Q&A groups, according to findings obtained. Secondly, knowledge Q&A groups, unlike other forms of media, foster a user-centered data flow environment [16], in which users gather data or knowledge from content providers largely through researching, inquiring, and connecting[17]. Users can also learn more about their interests by participating in a back-and-forth Q&A with further viewers or content providers in the forum under the postings they're involved in [18]. Operators can also express their problems and queries in other groups to get responses.

C. Emotional Cognition Theory (ECT)

It is a component of the emotional theory that is inspired by sociology. According to the hypothesis, once the body's data from outside commences the perception scheme, it'll be systematized and assembled by the perception scheme on the one side, as well as the data will activate the mind's positive or negative emotional feelings, which will then trigger some behavior propensities on either hand. In earlier studies, the extent of knowledge transfer[19] and social exchange theory[20], social capital theory[21], and other theories or approaches were used to analyze operator data exchange in the knowledge Q&A community. Knowledge discovery and critical thought are attributed by some experts to communities' users' information behavior. Individuals, on the other hand, are emotional beings who are easily influenced by their surroundings. The material they observe and their own emotions influence how they disseminate information. Furthermore, in the procedure of information analysis, the user's cognitive and emotions will perform a significant part in the information processing[22]. Operators in the group will display clear expressive inclinations when post due to the body's observation, that is in line with the main features of the ECT. As a result, this study integrates ECT with communal data dissemination behavioral science, which somewhat broadens the context of the theory's implementation but likewise aims to comprehend the emotive transformation of operators in the knowledge Q&A community throughout the pandemic by examining their temporal and spatial sentimental progression.

1.2 Approaches to OL Detection

A. Measures Based on Network Structure

This segment discusses methods for evaluating headship based on the network's structure, with topological variables serving as markers for influential nodes. Because the societal place is a significant aspect in OL heuristics based on vertex level, network pathways, or communal composition are employed to recognize possible leaders by their positioning inside the system. Figure 1 shows the different approaches to OL detection.

(a) Centrality measures

Calculating a vertex's centrality is a simple way of determining its importance. The centrality of vertices is a measurement of its significance in the network from a particular viewpoint[23] and can be stated in a variability of ways, resulting in a variation of centrality measures. The centrality measure is appropriate in a given situation and is determined by which feature of the network topology the metric reflects. The degree of the vertex, often known as centrality measures, is the most basic centrality metric. The more interconnected a vertex is, the more relevant it is in the network, according to this metric. From the standpoint of opinion leaders, it appears obvious that those with more connections have more opportunities for communication and impact. Eigenvector centrality is an initiative to convey the qualitative feature of a vertex's connections. The metric additionally consider the neighbors' centrality, based on the notion that links to more prominent vertices are more relevant than linkages to less prominent vertices. PageRank, which measures the significance of web pages in a search engine, is similar to eigenvector centrality[24].

(b) Community-driven measures

Several methods for estimating an individual's significance in a network exist, but most of them—for example, the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

standard centrality measures —do not consider the channel's community composition. A community is defined as a region of the network with a greater connection density and is commonly created by vertex that shares mutual traits[25]: persons, who regularly interact with one other, or clusters of compatible persons. Both internal and external connections are essential in terms of social position as an influence for OL. The positioning of people within a network's communal structure is an essential attribute, and it seems appropriate to consider their significance while judging their relevance. Embeddedness is one of the most basic metrics for identifying relevant vertices in a channel's community composition. It's the ratio of the overall degree to the internal degree. This metric measures how closely a vertex is linked to its community; yet, many vertices in network systems have no links to other societies and have strong integrated theoretical values [26]. Embeddedness' capacity to discriminate between significant and irrelevant vertices is harmed in such instances.

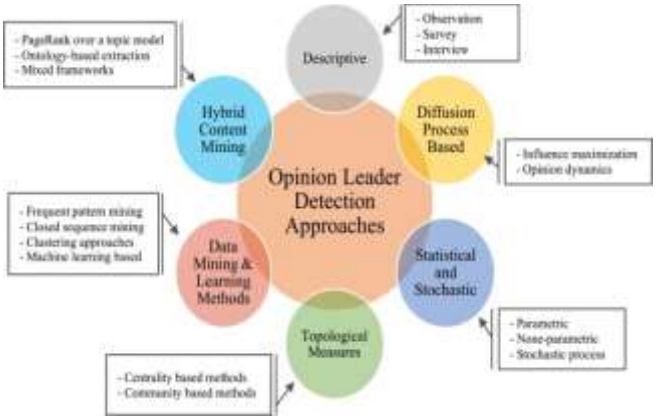


Figure 1- Different approaches to OL detection[27].

B. Methods Based on Interaction

The connections between persons at a given point in time are captured by overall network information. Interaction information, on the other hand, records dynamic interactions between persons, often taking into account the liberal dimensions of these interactions. Monitoring user communications can assist in determining the movement of influence in a network, which provides valuable information about the network's most influential individuals. The flow of impact in a system is strongly responsible for the information diffusion mechanism in the way that if the dispersion procedure is observed or demonstrated properly, the local or global impact of a node may be deduced in general. All of the techniques in this area rely on interaction data, and the majority of them also examine structure network data; the algorithms vary from greedy search tactics to mining algorithms or clustering.

C. Methods Based on Content Mining

This type of strategy is based on network construction and data gathered from the content that numerous nodes in the network expose and share. Subject information (because OL is topic reliant on), explicit user attributes, and sentiment are all extracted from the text. Although not all tactics are created expressly for identifying leaders, the approaches can be used in this situation.

D. Methods Based on Content and Interaction

To find OL, the techniques in this area use both interaction information and content, as well as a network structure. The approaches take advantage of temporal cues including feedback flow, and specific data, as well as exact composition, retrieved such as subject recognition and user-related aspects. Nearly methods in this area take into account network topology, whereas others rely simply on content and temporal data[28].

1.3 A general methodology for OL selection using Ant Colony Optimization (ACO) approach-

OLs are individuals in an SN who have the most impact on the acceptance or enjoyment of products/services by other persons. Formally and informally, OL controls the distribution and adoption of a new item, as well as the decisions of others through diverse communication networks. Consequently, corporations' interest in OL has improved, making them a required primary marketing objective. Although there are so many distinct sorts of OL around the globe, selecting the appropriate one can be challenging. Several criteria will be used to identify an OL. These qualities were determined by reviewing the relevant literature. Due to the presence of diverse leaders, the feature set of these leaders will be expansive, and there may be redundant or irrelevant characteristics that impact selection. These were screened to eliminate undesirable characteristics.

A. Feature-based selection

Characteristics selection is performed to eliminate incorrect features to limit the number of OL. The dimensionality curse is caused by the fact that a diverse set of features makes processing difficult and exacerbates the complexity of computations. The features-based selection strategy is used to pick a subset of the original set while preserving the precision of the original set. This is accomplished by deleting unnecessary characteristics from the collection and sorting out suitable features. Features, effectiveness, and scalability can all be enhanced. The choice of appropriate leaders is based on their distinct characteristics and is required when working with a large variety of leaders (characteristics) datasets. Real-world issues necessitate the use of feature-based selection due to the abundance of obnoxious, inappropriate, or deceptive elements. Based on these criterion features, the suggested method would identify suitable opinion leaders for assigning various organizational activities.

B. Ants Generation

ACO was suggested as a swarm intelligence technique. Ants have inspired a wide range of approaches and procedures, with ACO, a general optimization technique, being the most well-known and frequently employed. Various ant species' foraging behavior is utilized to enhance ant colonies. When ants migrate from one site to another, they deposit pheromones (chemicals) on the ground to guide other ants in the colony (members). Similarly, ACO solves optimization problems. Typically, the ACO technique is employed to determine the ideal solution. To solve difficulties, ants utilize two variables: heuristic knowledge and the frequency of pheromones. Artificial ants that converse with one another have the potential to provide superior results. The pheromone trail parameters are achieved by indirect interaction (sensing of the pheromone) between different ants. Ants do not transform themselves; rather, they affect

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

how other ants perceive and understand the situation. ACOs can be utilized to address a range of difficulties.

C. Opinion Leader Selection

The procedure of picking a leader begins with the formation of an AC, which travels along several paths (edges) and selects a leader based on the pheromone value of each edge. If stopping requirements for ant visits are met, the ants will leave (traverse) and the best subclass of OLs will be identified, which will then be utilized in a variety of marketing or industrial contexts. If the traversal of ants during the visit does not fulfill the halting conditions, the pheromone values are modified and the cycle is repeated. Figure 2 shows the general flow for opinion leader detection with the Ants colony optimization algorithm. The work starts with the ant's generation with OL selection. If the criteria of the selection meet, then the program may end. And if the criteria don't meet then the new values should be modified and initialized again and the process continues till it finds the best opinion leader.

D. Formation of Results

The selection of OL based on their characteristics begins with the creation of a random number of fake ants. On the graph, the number of ants resembles the number of linked OLs. The procedure of creating a graph from a single node is initiated by each ant's traverse. Ant moves probabilistically from a starting place, traversing multiple nodes before process completion. The resulting group of OLs is collected and examined to determine the ideal subset. When the best OLs have been identified and the results are made public.

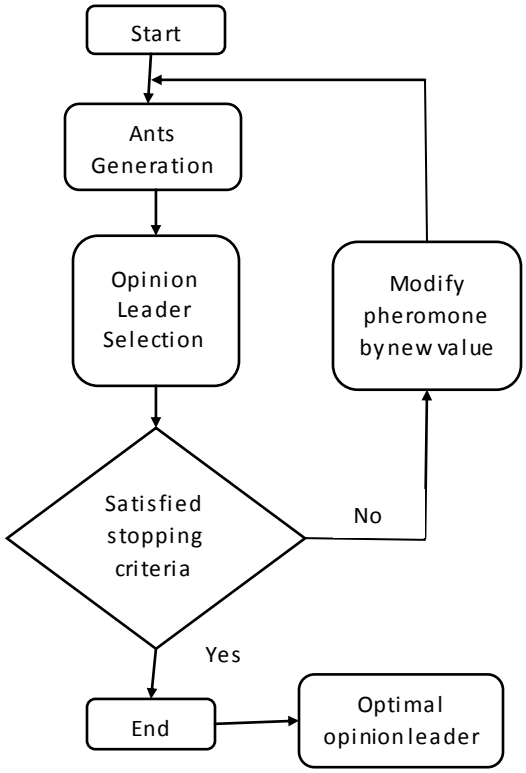


Figure 2 – A general framework for opinion leader detection[29].

II. RELATED WORK

Ma and Liu[13] analyzed the characteristics of 3 seed networks and identified OL in the Fukushima nuclear crisis using the Superedge Rank algorithm. Jiang et al. [30] used MapReduce to develop and deploy a BBS OL mining solution founded on an adaptive PageRank algorithm.[31] took the main procedure of Internet browsing–the PageRank model–and proposed an approach to determine Internet opinion leaders by integrating the understanding of the effect of linguistic data and emotional preferences; they also tested the strategy with an empirical investigation. Cheng et al. [31] used the IS Rank method, which integrates the PR values of the PageRank algorithm with appreciation degree, to merge impact with SA depending on the content of postings and filtered OL. Deng et al.[32] proposed a node closeness approximate calculation method for discovering core OL on Micro Blog graphs, which offers increased accurateness and moves better, and created a Micro Blog crawling and analytics tool based on SINA Micro Blog APIs. PageRank is an efficient sorting algorithm, but its performance degrades dramatically as the number of available nodes grows. Jing and Lizhen[33] presented a hybrid data mining technique depending on user characteristics and conversation networks that consists of three parts: a technique for analyzing users' authority, activity, and impacts, a method for considering sentiment polarity in an interaction network, and a consolidated approach for recognizing microblog OL based on the HITS algorithm.

OLs on Twitter usually have a large following and have a significant influence on their followers. Their blog is usually full of useful data and breaking news for everyone else[34]. Followers enjoy sharing their opinion leaders' tweets data because they believe it is important for others to know. All of these qualities of Twitter OL reinforce the notion that knowing who the OLs are on Twitter is valuable. Existing options, on the other hand, do not focus on specific opinion leaders, but rather on general ones, as we have described. Hema Banati and Monika Bajaj[35] compare the effectiveness of evolutionary algorithms to the current greedy technique for this application. It evaluates the appropriateness of two popular EA's, Differential Evolution (DE) and Firefly (FA), to the k-Max-Influence problem. Epinions, Wiki-Vote, Slashdot datasets were used in the research. In terms of maximal influence experienced as well as advantage obtained by boosting the value of k, the results demonstrated that both approaches DE and FA outperform the Greedy strategy. FA consistently outperforms DE among evolutionary techniques. The data demonstrate that FA keeps its outcomes consistent and has a larger chance of scoring than DE and Greedy.

2.1 Research Gaps

When discussing the application to real-world circumstances, the following potential issues and restrictions must be considered:

1. Availability of information. This feature may affect the accessibility of necessary data. Researchers utilize solely publicly available, anonymized user data. So certain data might not be available.

2. Enormous quantities of data.

3. Information that is poorly organized and contains a great deal of noise Therefore, information should first be organized and then examined. The last two characteristics necessitate specialized data collecting, processing, etc.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

approaches, hence necessitating the employment of qualified specialists.

Table 1- Review of earlier proposed approaches.

Author	Method	Evaluation Results	Ref.
Junwei Wang and Dingwei Wang [2008]	Particle swarm optimization	Goose team optimization attempts to incorporate several animal aggregating patterns into the optimization process. The program can effectively optimize four difficult nonlinear functions thanks to the goose team flight technique.	[36]
W.R.M. U.K. Wickramasinghe and X. Li [2008]	Differential Evolution	MDEPSO is a simple and reliable EMO method that performs similarly to NSGA-II.	[37]
Yiyi Zhao et al. [2018]	confidence-based opinion dynamics model	The trust factor that opinion followers have for OL has a superior effect on their ability to influence others.	[38]
sMichał Weskida et al. [2018]	Evolutionary Algorithm parameters	They found that the sets of ranking parameters employed in the evolutionary process had a strong correlation, often around 0.9.	[39]
M.R. Martínez-Torresa et al. [2018]	Particle Swarm Optimization (PSO)	The findings can aid in a better understanding of the authentication mechanism in eWOM groups and the avoidance of reputation manipulations by fraudulent accounts.	[40]
Lokesh Jain and Rahul Katarya [2018]	Firefly Algorithm	The suggested methodology is superior and ideal for identifying the opinion leader in a community as well as across the SN at large.	[41]
Sofya A. Semenkovich et al. [2019]	SVM	When comparing the proposed algorithm to current ones, it was discovered that all of them consider certain nodes to be prospective leaders.	[42]
Lokesh Jain et al. [2019]	whale optimization algorithms	Depending on factors and computational time, they also concluded that the communal partition approach is even superior to the other communities' detection techniques.	[43]
S. Mohammad Aghadam et al. [2020]	Grey Wolf Optimizer (GWO) algorithm	The number of known OL discovered by this algorithm is large, and the proposed method has the advantage of being compatible with a variety of criteria and producing long-term results in a variety of methods.	[44]
Afsaneh Razghandi and Yasser Elmi [2020]	Leach algorithm	With a correlation value larger than 0.7, the algorithm has a considerable association with the classic centrality indexes.	[45]
Dr.K.Saraswath et al. [2020]	Multi-Objective function based Ant Colony Optimization (MOFACO)	The classification performance of IG-SVM RBF (100,0.1) outperforms IG-CART by 17.26%, according to the results.	[46]
Xu Qiang et al. [2021]	Ant Colony Optimization	This research contributes significantly by helping in the selection of an appropriate OL based on the criterial qualities.	[29]

As social media and activities evolve, online SN has become more integrated into people's work and daily lives. OLs are activists who use media to spread information and filter statements in mass media. As a result, effective monitoring of OL may impact the dissemination and evolution of public sentiment to some level. Several traditional techniques of OL mining rely primarily on the user's network structure, ignoring the importance and relevance of contextual information in the development of opinion leaders. Furthermore, these methodologies only rank users' impact on a global scale and are ineffective at discovering low-influence local opinion leaders.

Despite these important findings, properly comprehending opinion dynamics via social processes is never an easy task.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

Individual information obtained from a variety of avenues is an important component in decision-making and opinion formation. It lays the groundwork for forecasting the outcomes of various courses of action. However, not everyone has the same knowledge foundation. When information is difficult to absorb or obtain owing to incompleteness, people often shape their beliefs by interacting with others who share their viewpoints on the same subjects. A basic example of this is that it is easy for someone to generate views on places they have never seen or foods they have never tried because decisions can be formed based on information obtained from others. Popular computational approaches have been concentrating on how opinion emerges in a social setting for a few decades, in which opinion is either condensed to an arbitrary number or analyzed in relevant to a particular situation.

2.2 Discussion

People's daily life have become increasingly reliant on Web 2.0. Several virtual communities, such as blogs, and websites, are growing more engaged with time. People can readily voice their views in public and observe what others have to say. The initial difficulty in identifying OLs is determining how a leader is characterized in various circumstances. If we define thought leaders as persons who can greatly influence others, then leadership can be regarded from a variety of viewpoints. A leader's potential to convey persuasion is intuitively represented in the structural qualities of the SN. A typical definition of the leader notion is the origin of an action that spreads further and quicker than other activities. By integrating structural features, influence flow, and diffusion processes in networks, a comprehensive solution can be attained. Essential to determine OL is their inherent qualities, which, in conjunction with other variables, determine their standing. By evaluating personal attributes and the information generated by leaders, including online posts, blog entries, comments on news stories, and even endorsements of others' work through sharing, one can indirectly analyze these intrinsic qualities. In actuality, all of these factors contribute to OL status, and ideally, methodologies for identifying OLs would take into account all of these factors. Nevertheless, it is challenging to build broadly applicable procedures that take into account all of these issues. Consequently, the approach that is most suitable to a particular problem relies on its environment.

Applications

When OLs are addressed, information can be conveyed more successfully in a variety of areas. In the medical industry, for instance, identifying OLs has been utilized to promote novel and effective therapies, such as HIV prevention and child health development. To successfully communicate information, climate change awareness efforts have utilized opinion leaders. In the middle to late years of the first decade of this century, Al Gore coordinated two initiatives that selected OL to give demonstrations and educate the general public on climate change issues. Detection of OLs is especially important in advertising, promotion, and product adoption, where organizations want to attract potential customers to their products as efficiently as feasible. Political campaigns can also profit from recognizing and contacting OL, as seen by George W. Bush's 2004 presidential election campaign, in which OL were chosen to endorse the campaign. Significant who the OL are can be advantageous in several different areas, making their accurate and efficient recognition a crucial responsibility.

III. CONCLUSION

Nowadays, networks are the most popular medium among the general public, and practically everyone is connected to them, resulting in a slew of societal difficulties as virtual networks spill over into the real world. OLs play a critical role in the distribution of information on a variety of topics. Previous studies used quantitative clustering approaches to recognize OL or used social assessment approaches based on interview self-reports or reviews. Some researchers have looked at online posts to establish OL's social tendencies and interactions with their followers. As a result, this research must suggest a way of analyzing the features of social platform posts, and their interactions. To establish whether or not the dissemination of innovation succeeds, it is important to identify OL for an opinion leadership intervention. Few analyses have provided a comprehensive framework for researchers to use in determining opinion leaders. The goal of the current study was to present a synthesis of the methodologies for identifying OL using SN analysis techniques, based on this need. This research combined communication network data gathering methods with social network analysis to determine opinion leaders, and it presented how SN data were obtained in two different social systems. In social network analysis, three metrics of network centrality are commonly used: degree, betweenness, and closeness. It will be exciting to see if other amalgamation approaches can be used to generate leaders in addition to Differential Evolution. We'd also like to investigate if this hybrid model can address rotational multi-objective situations, which are difficult to solve with current EMO methods.

REFERENCES

[1] T. Chen, J. Shi, J. Yang, G. Cong, and G. Li, "Modeling Public Opinion Polarization in Group Behavior by Integrating SIRS-Based Information Diffusion Process," *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/4791527.

[2] M. G. J. Ramakrishnan, D. Mavaluru, K. Srinivasan, A. Mubarakali, C. Narmatha, "Opinion Mining using Machine Learning Approaches: A Critical Study," *Vellore Inst. Technol.*, 2020, [Online]. Available: https://research.vit.ac.in/publication/opinion-mining-using-machine-learning-approaches-a-critical-study.

[3] W. Haoxiang and S. S, "Big Data Analysis and Perturbation using Data Mining Algorithm," *J. Soft Comput. Paradig.*, vol. 3, no. 1, pp. 19–28, 2021, doi: 10.36548/jscp.2021.1.003.

[4] S. S. Lokesh Jain, Rahul Katarya, "Recognition of opinion leaders coalitions in online social network using game theory," *Researcher*, 2020, [Online]. Available: https://www.researcher-app.com/paper/5267398.

[5] I. Chien, C. Y. Lin, and I. H. Wang, "Community detection in hypergraphs: Optimal statistical limit and efficient algorithms," *Int. Conf. Artif. Intell. Stat. AISTATS 2018*, vol. 84, pp. 871–879, 2018.

[6] B. Cai, L. Zeng, Y. Wang, H. Li, and Y. Hu, "Community detection method based on node density, degree centrality, and k-means clustering in complex network," *Entropy*, vol. 21, no. 12, 2019, doi: 10.3390/e21121145.

[7] J. Cao, "Compactness Preserving Community Computation Via a Network Generative Process," *IEEE*, 2020, [Online]. Available: https://ieeexplore.ieee.org/document/9548676.

[8] H. L. Zhan Bu, "Graph K-means Based on Leader Identification, Dynamic Game, and Opinion Dynamics," *Semant. Sch.*, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

[9] N. Meißner, "Opinion leaders as intermediaries in audience building for independent films in the Internet age," *SAGE Journals*, 2014.

[10] Boatwright, "Exploring Online Opinion Leadership: An Analysis of the Influential Users on Twitter During the Online Conversation Around Anthem Protests by Prominent Athletes," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.

[11] Y. Xiong, Z. Cheng, E. Liang, and Y. Wu, "Accumulation mechanism of opinion leaders' social interaction ties in virtual communities: Empirical evidence from China," *Comput. Human Behav.*, vol. 82, pp. 81–93, 2018, doi: 10.1016/j.chb.2018.01.005.

[12] W. Zhang and H. He, "Identifying and evaluating the internet opinion leader community based on k-clique clustering," *Acedemia*, 2013, [Online]. Available: https://www.academia.edu/37514900/Identifying_and_evaluating_the_internet_opinion_leader_community_based_on_k_clique_clustering.

[13] N. Ma and Y. Liu, "SuperedgeRank algorithm and its application in identifying opinion leader of online public opinion supernetwork," *Expert Syst. Appl.*, vol. 41, no. 4 PART 1, pp. 1357–1368, 2014, doi: 10.1016/j.eswa.2013.08.033.

[14] B. J. J. Zhe Liu, "PennState," *Identifying Predict. desire to Help Soc. Quest. answering*, 2017.

[15] C. S. Sanghee Oh, J. Oh, "The use of information sources by internet users in answering questions," *Semant. Sch.*, 2008, [Online]. Available: https://www.semanticscholar.org/paper/The-use-of-information-sources-by-internet-users-in-Oh-Oh/3ece01da3162f9468e3a31a1e2a032b702461321.

[16] Y. Chen, T. Dong, Q. Ban, and Y. Li, "What concerns consumers about hypertension? A comparison between the online health community and the q&a forum," *Int. J. Comput. Intell. Syst.*, vol. 14, no. 1, pp. 734–743, 2021, doi: 10.2991/ijcis.d.210203.002.

[17] 693–702 Gazan, R. Microcollaborations in a social Q&A community. Inf. Process. Manag. 2010, 46, "Microcollaborations in a social Q&A community," *Inf. Process. Manag.*, vol. 46, no. 6, pp. 693–702, 2010, doi: 10.1016/j.ipm.2009.10.007.

[18] M. L. G. Zhenya Tang, Leida Chen, "How to keep brand fan page followers? The lens of person-environment fit theory," *emarland*, 2018.

[19] T. Guan, L. Wang, J. Jin, and X. Song, "Knowledge contribution behavior in online Q&A communities: An empirical investigation," *Comput. Human Behav.*, vol. 81, pp. 137–147, 2018, doi: 10.1016/j.chb.2017.12.023.

[20] J. Jin, "Why users contribute knowledge to online communities," *ACM diigital Libr.*, 2015.

[21] C. A. Johnson, "Social capital and the search for information: Examining the role of social capital in information seeking behavior in Mongolia," *J. Am. Soc. Inf. Sci. Technol.*, vol. 58, no. 6, pp. 883–894, 2007, doi: 10.1002/asi.20561.

[22] Jonah Berger, "What Makes Online Content Viral?," *SAGE Journals*, 2012.

[23] L. SergeyBrin, "The anatomy of a large-scale hypertextual Web search engine," *sciencedirect*, 2000.

[24] L. C. Freeman, "A Set of Measures of Centrality Based on Betweenness," *JSTOR*.

[25] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, no. 3–5, pp. 75–174, 2010, doi: 10.1016/j.physrep.2009.11.002.

[26] A. Lancichinetti, J. Saramäki, M. Kivelä, and S. Fortunato, "Characterizing the community structure of complex networks," *PLoS One*, vol. 5, no. 8, pp. 1–8, 2010, doi: 10.1371/journal.pone.0011976.

[27] S. M. HosseiniBamakanaIldarNurgalievabQiangQ, "Opinion leader detection: A methodological review," 2019, [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417418304950.

[28] P. Parau, C. Lemnaru, M. Dinsoreanu, and R. Potolea, *Opinion Leader Detection*, no. 1. Elsevier Inc., 2017.

[29] X. Qiang, Z. Huiqi, F. Ali, and S. Nazir, "Criterial Based Opinion Leader's Selection for Decision-Making Using Ant Colony Optimization," *Sci. Program.*, vol. 2021, 2021, doi: 10.1155/2021/4624334.

[30] L. Jiang, "BBS opinion leader mining based on an improved PageRank algorithm using MapReduce," *Res. Gate*, 2013.

[31] C. C. Huang, W. Y. Liang, P. A. Chen, and Y. C. Chan, "Identification of opinion leaders and followers—a case study of green energy and low carbons," *Appl. Sci.*, vol. 10, no. 23, pp. 1–16, 2020, doi: 10.3390/app10238416.

[32] X. Deng, Y. Li, and S. Lin, "Parallel Micro Blog Crawler Construction for Effective Opinion Leader Approximation," *AASRI Procedia*, vol. 5, pp. 170–176, 2013, doi: 10.1016/j.aasri.2013.10.074.

[33] L. Jing, "Identification of Microblog Opinion Leader Based on User Feature and Interaction Network," *ACM diigital Libr.*

[34] M. Ghosh, K. Gupta, and S. Susan, "Aspect-Based Unsupervised Negative Sentiment Analysis," *Lect. Notes Data Eng. Commun. Technol.*, vol. 57, no. February, pp. 335–344, 2021, doi: 10.1007/978-981-15-9509-7_29.

[35] H. Banati and M. Bajaj, "Evolutionary Approach for k-Max Influence Problem," vol. 1, pp. 163–172, 2013.

[36] J. Wang and D. Wang, "Particle swarm optimization with a leader and followers," *Prog. Nat. Sci.*, vol. 18, no. 11, pp. 1437–1443, 2008, doi: 10.1016/j.pnsc.2008.03.029.

[37] W. R. M. U. K. Wickramasinghe and X. Li, "Choosing leaders for multi-objective PSO algorithms using differential evolution," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5361 LNAI, pp. 249–258, 2008, doi: 10.1007/978-3-540-89694-4_26.

[38] Y. Zhao, G. Kou, Y. Peng, and Y. Chen, "Understanding influence power of opinion leaders in e-commerce networks: An opinion dynamics theory perspective," *Inf. Sci. (Ny).*, vol. 426, pp. 131–147, 2018, doi: 10.1016/j.ins.2017.10.031.

[39] M. Weskida and R. Michalski, "Finding influentials in social networks using evolutionary algorithm," *J. Comput. Sci.*, vol. 31, pp. 77–85, 2019, doi: 10.1016/j.jocs.2018.12.010.

[40] M. R. Martínez-Torres, F. J. Arenas-Marquez, M. Olmedilla, and S. L. Toral, "Identifying the features of reputable users in eWOM communities by using Particle Swarm Optimization," *Technol. Forecast. Soc. Change*, vol. 133, no. April, pp. 220–228, 2018, doi: 10.1016/j.techfore.2018.04.017.

[41] L. Jain and R. Katarya, "Discover opinion leader in online social network using firefly algorithm," *Expert Syst. Appl.*, vol. 122, pp. 1–15, 2019, doi: 10.1016/j.eswa.2018.12.043.

[42] S. A. Semenkovich and O. A. Tsukanova, "On the Algorithms of Identifying Opinion Leaders in Social Networks," *Procedia Comput. Sci.*, vol. 162, no. Itqm 2019, pp. 778–785, 2019, doi: 10.1016/j.procs.2019.12.050.

[43] L. Jain, R. Katarya, and S. Sachdeva, "Opinion leader detection using whale optimization algorithm in online social network," *Expert Syst. Appl.*, vol. 142, p. 113016, 2020, doi: 10.1016/j.eswa.2019.113016.

[44] S. M. Aghadam, F. S. Gharehchopogh, and M. Masdari, "Opinion Leaders Selection with Grey Wolf Optimizer Algorithm on," vol. 13, no. 2, 2021.

[45] L. C. C. By, S. Date, P. Date, C. Elmi, and I. Leach, "Improved Leach algorithm for detecting opinion leaders in social networks," 2020, doi: 10.36227/techrxiv.13474641.v1.

[46] K. Saraswathi, N. T. Renukadevi, and S. Karunakaran, "Multi ant colony optimization for opinion classification," *Int. J. Sci. Technol. Res.*, vol. 9, no. 3, pp. 4554–4560, 2020.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:59 UTC from IEEE Xplore. Restrictions apply.

Check for updates

Received: 3 March 2022 | Revised: 24 April 2022 | Accepted: 9 May 2022

DOI: 10.1002/app.52816

REVIEW

Applied Polymer SCIENCE WILEY

A review on self-healing polymers for applications in spacecraft and construction of roads

Megha Goyal[1] | **Satya Narayan Agarwal**[2] | **Nitu Bhatnagar**[1]

[1]Department of Chemistry, Manipal University Jaipur, Jaipur, India

[2]Department of Electrical Engineering, Manipal University Jaipur, Jaipur, India

Correspondence
Nitu Bhatnagar, Department of Chemistry, Manipal University Jaipur, Jaipur, India.
Email: niturbhatnagar@gmail.com

Abstract

Self-healing polymers (SHP) are the polymers that have the capability to recover the material from any physical damage by responding to the damages or cracks in the material and heal the same. This self-healing technology gives the polymers ability to recover from the cracks at an early stage by preventing them from any catastrophic failures, which increases the scope of applications of these materials. This review article highlights the key points on extrinsic self-healing polymer and intrinsic self-healing polymer by comparing their healing efficiency, advantages, disadvantages, and challenges in the prospect of their future development as well as their possible applications in space industry and construction of roads.

KEYWORDS

applications, composites, friction, nanoparticles, nanowires and nanocrystals, wear and lubrication

1 | INTRODUCTION

Self-healing polymers (which can heal or repair by itself) have been studied for their interesting applications in space industries, robotics, construction materials, etc.

The concept of Self-healing polymers started as welding polymers in 1950 by interfacial molecular diffusion. Further in 1952, Bueche et al. gave the correlation between volume viscosity and self-diffusion constant of $(C_8H_8)_n$ and Butyl ester of acrylic acid.[1] Then, in 1970s, it was revealed that the self-healing properties of poly(1-acetyloxiethylene) increases in strength during the period of relaxation which was exactly related to the healing temperature.[2,3] In the year of 1980, PMMA (polymethyl-methacrylate) was improved in self-healing efficiency during the time that temperature reached to the T_g (glass transition temperature) because of the accelerated interface diffusion.[4] Later, Wool and O'connor[5] discussed five phases of self-healing polymers as (Figure 1).

- polymer chains were rearranged on two bisecting surfaces,
- surface approach done by bringing the damaged surface together,
- wetting is the third stage where both surfaces have to wet each other,
- diffusion should be limited for a small area, and
- randomization is the last damage surface where it healed the material after diffusion.[5]

These polymers get healed by physical or chemical phenomenon.[6] Physical self-healing phenomenon involves inter-chain diffusion,[5] phase-separated morphologies,[7,8] shape-memory effects,[9] or involve superparamagnetic nanoparticles,[10] while chemical self-healing phenomenon involves incorporation of covalent,[11–13] free-radical,[14,15] or supramolecular dynamic bonds.[16–18] Self-healing is also possible by combination of both physical and chemical phenomena which involves van der Walls interaction,[19] encapsulating microcapsules,[20] and motion of cardiovascular networks.[21] Structural damage occurs at micro or macro level. At macro level, damage breaks the material while damage at micro-level affects the material characteristics.[22]

 © 2022 Wiley Periodicals LLC.

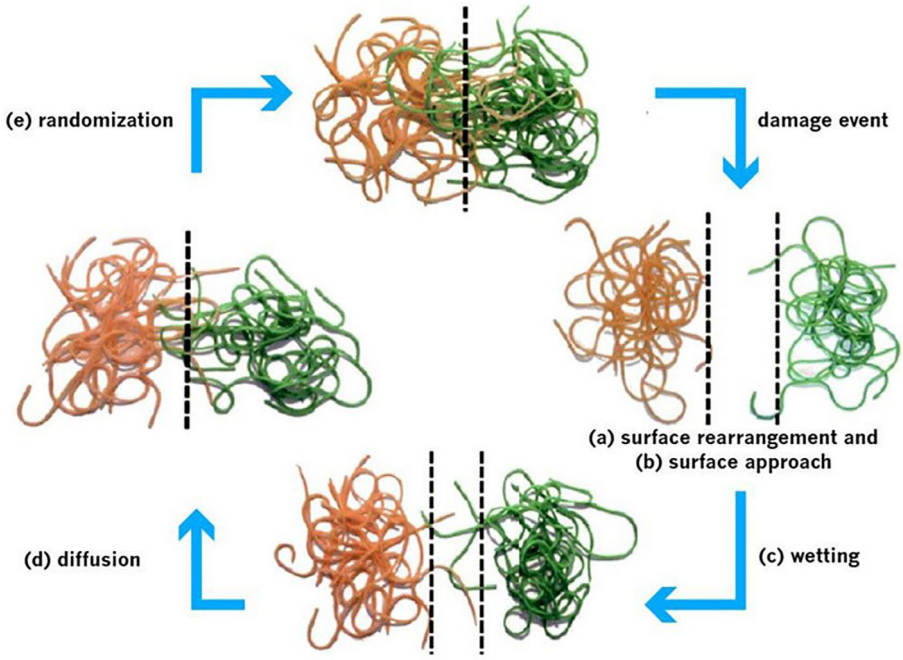


FIGURE 1 Five phases of crack healing representation: (a) rearrangement, (b) surface approach, (c) wetting, (d) diffusion, (e) randomization[5] [Color figure can be viewed at wileyonlinelibrary.com]

Recently, a lot of progress has been done for the development of new materials which can be used to fabricate damages that can takes place in structural materials after their prolonged use.

Like, in the application of concrete, cracks can happen at any time due to the weakness of material. Fatigue load decreases the concrete toughness which increases the permeability and lead to the decrement of durability, and lifespan of construction material.[23] In medical application, the major defects shown in the material distortion and the excess usage which leads to the mechanical failure in biomedical implants.[24] In the application of coating system, Surface scratching and microcracking can stimulate the creation of defects, which can lead to macroscopic damage and losses its properties. High packing efficiency, low mass, broad volume, and launch cost are reason for application in space materials.[25] These kinds of damage initiation might lead to more significant damage or perhaps catastrophic failure. These catastrophic failures might have a substantial influence on current and new structural repair, maintenance, and rehabilitation. As a result, this catastrophic failure might have major economic, social, and environmental consequences and preventative measures should be taken at the microlevel to avoid significant harm.

Hence, recent literature[22,26] has come up with the importance of self-healing polymers as an alternative for protecting structural parts from damage, which has motivated us to undertake this topic of research. This review paper gives an idea about the self-healing polymers and the healing mechanism involved in aerospace and construction applications. It discusses the principle involved in the interaction mechanism of self-healing polymers and their effectiveness. In the next section, the advantages and disadvantages of the self-healing polymers in the space and construction applications, and the current challenges of self-healing technologies are also discussed. Thus, this review paper gives a broader idea of self-healing polymers, which leads to improve the life cycle of materials.

2 | CLASSIFICATION OF SELF-HEALING POLYMERS

Self-healing polymers have the capability to deal with any damage efficiently.[27] Paul et al.[22] has discussed that at micro level, damage can be self-healed by three-step process: triggering, traveling, and repairing. Triggering is the first step that shows the need to react immediately to the damaged area by the healing medium. Second step is traveling in which healing travels to the damaged area while repairing is the third step, where the chemical restore the damage. Based on the mechanism and working chemistry, self-healing polymers (SHP) are classified as shown in Figure 2.[28]

Self-healing is achieved by the different mechanism which are introduced to repair the damage. Based on this ability they are classified into intrinsic and extrinsic.[29] Intrinsic self-healing polymers undergo dynamic covalent chemistry which includes Diels–Alder reaction, dynamic urea bond, imine bonds, and radical exchange and non-covalent chemistry includes H-bonding, ligand–metal bonding, and host–guest interaction (Figure 3).[28] Intrinsic

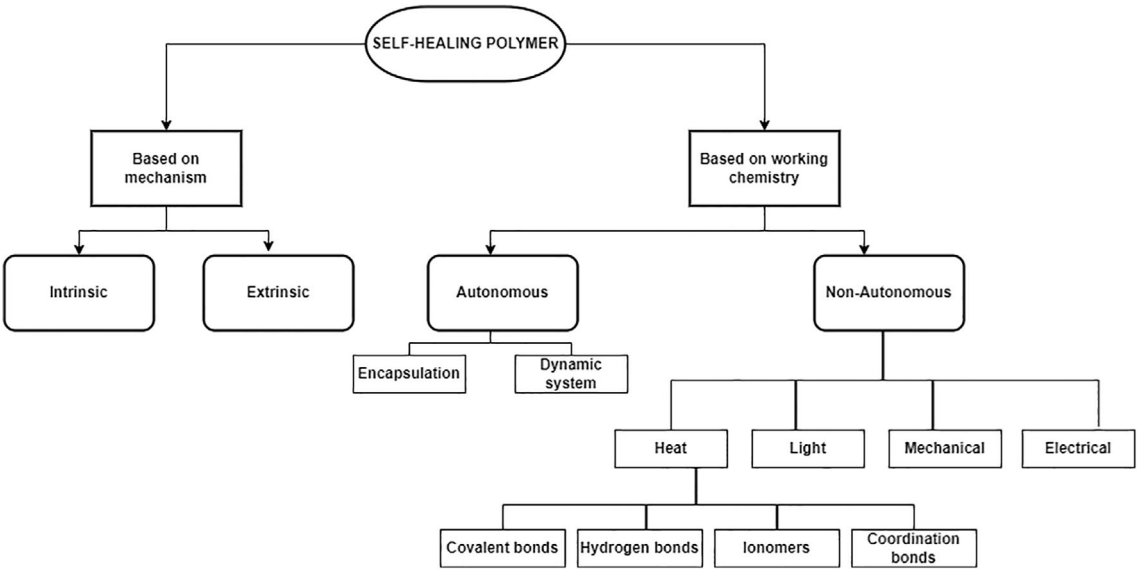


FIGURE 2 Classification of self-healing polymer[28]

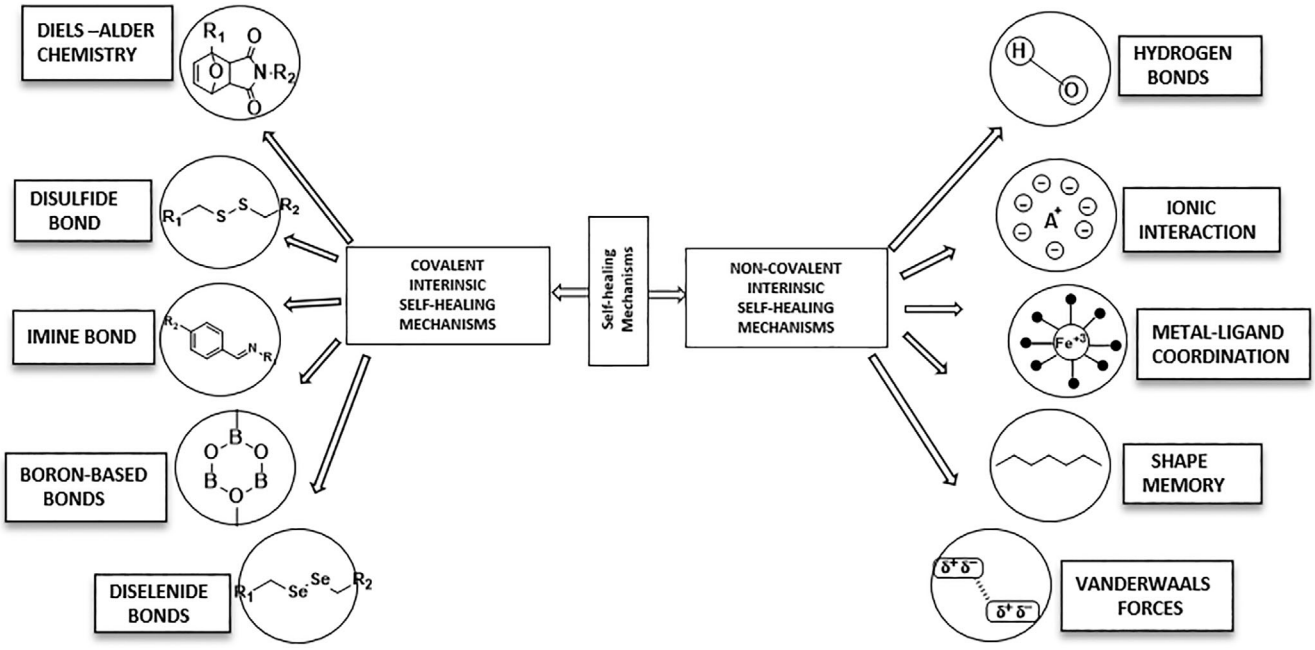


FIGURE 3 Covalent and non-covalent intrinsic mechanism[29]

self-healing is triggered external stimuli like temperature, pH, light, pressure, and others. The healing can be done for multiple cycles due to which it does not need any catalyst. The only disadvantage of intrinsic self-healing is that it is restricted to thermoplastic materials like acrylic, polyester etc., which lowers down its strength and make the bond weaker.[28]

In contrast, extrinsic self-healing is based on an external healing which is dispersed in capsules or vascular networks. White and his co-workers designed the extrinsic self-healing polymer by using liquid healing agent in microcapsules.[20] Its initiation can be triggered by external/internal damage both. It can be dispersed into the matrix with the process of encapsulation techniques. The disadvantage of extrinsic healing is that it is limited to one-time process and involves complicated factors such as resin viscosity etc.[28] A brief discussion on the same has been highlighted in Table 1.[30] Table 2 gives the brief difference between intrinsic and extrinsic mechanism in detail.

Depending on the working chemistry, self-healing polymers are categorized into (i) autonomous self-healing polymers, (ii) non-autonomous self-healing polymers.

TABLE 1 For extrinsic self-healing mechanism

Self-healing system	Category	Advantages	Disadvantages	Applications
Capsule based system	Catalyst	Catalytic and epoxy filled capsules are included Without advanced techniques, capsules for epoxy load can be identified	One time healing cycle can be done Grubb's catalyst provides best result as it required high cost and sensitive to environmental condition such as moisture and oxygen	Applicable at aircraft, aerospace, coatings, paintings etc
	Dual capsule	There is no catalyst Matrix capsules may readily be dispersed	One time healing cycle can be done Due to its hardness, it is difficult to be encapsulated	Anti-corrosion coating for aerospace, automobile, etc Helpful to avoid cracking during healing and to cure crack
	Mono-capsule	Requires a single type of healing capsule Dispersed easily into the matrix	Mostly, solvents are toxic	Enhances the sports equipment Any prosthesis implant body that cannot be maintained, for example, limbs, joints, etc
Vascular based system	One-dimensional	It is possible to conduct many healing	During the filling and release of healing agents, core fiber blockages were caused by viscosity and system diameter	Electric circuit that heals and restores after damage conductivity
	2D and 3D	Multiple healing can be carried out and all the directions of the healing distributed	During the filling and release of healing agents, core fiber blockages were caused by viscosity and system diameter	In order to prevent harm, good packaging is necessary. SMP is very handy for packaging food

Autonomous self-healing polymers are those polymers which can recover their shape automatically or which can respond to the damage and get recovered. Carolyn Dry reported the first autonomic healing polymer.[31] Non-autonomous self-healing polymers need the stimuli to heal the damaged surface. Suman Thakur[28] has given a comparison of Autonomous and Non-Autonomous self-healing polymer which is as shown in Table 3.

3 | CHEMISTRY OF SELF-HEALING POLYMERS AND ITS MECHANISM

The qualities of self-healing polymers belong to the type of smart material that can repair internal defects in any matrix and restore mechanical properties like tensile strength and also extend the time period of the materials, and improve performance while using them for various application, that is, construction, medical, aerospace, defense, etc. Understanding and engineering effective polymerization processes is required for the core notion of a polymer's self-repairing mechanism. To give the properties of synthetic self-healing polymers, many approaches have been devised which includes subjecting them to the chemical, physical, or thermal stimuli. Hence, the healing effectiveness of a polymer is also determined by its high chain mobility and glass transition temperature (T_g). The toughness ratio between the pristine and healed samples is commonly used to measure a material's healing effectiveness. The next sections give explanations of various methods and mechanisms[32] which helps to provide synthetic polymers with self-healing properties.

3.1 | Photo-irradiation-assisted self-healing

Cheng and group described a photo-irradiation approach for resin-based self-healing.[33] The fragmented films of crosslinked polyurethanes ($C_{17}H_{16}N_2O_4$) were generated on glass substrates by reacting 1,3-bis(eugenyl) glycerol

TABLE 2 Shows a comparison between intrinsic and extrinsic mechanism

Parameters	Intrinsic mechanism	Extrinsic mechanism
Definition[23]	No external substance is needed to repair the damage	It required the external healing material in the capsule or in the vascular network
Properties[28]	Fast, self-healing can be possible Multiple healing can be possible No catalyst or healing agent should be needed to trigger healing	Autonomic- fast self-healing can be possible Multiple healing can be possible Easy encapsulation and simple techniques can be used
Mechanism[28]	It reform the chemical bonds	It requires the healing agent and the catalyst
Healing system[28]	It includes covalent, that is, Diels–Alder reaction, dynamic urea bond etc., and non-covalent chemistry, that is, hydrogen bond, metal–ligand bond etc	It includes an external healing agent mixed up in capsule and vascular networks

TABLE 3 Comparison between autonomous and non-autonomous self-healing polymer

Parameters	Autonomous	Non-autonomous
Definition	Repair can be healed by itself	The chemical is released and the damaged surface of a polymer is healed if an appropriate stimulus (such as heat or light) is applied
Properties	No catalyst or healing agent should be needed to trigger healing	Autonomic- fast self-healing can be possible Multiple healing can be possible Simple Techniques can be used
Mechanism	It is healing can be done under encapsulation or dynamic system	It requires the external stimuli such as heat, light, mechanical bond, and chemical bond

diether with hexamethylene diisocyanate. UV irradiation was employed to test the qualities of a self-healing polymer in this study. They showed that within 5 min, the damaged area of the polymer film self-healed, implying that generated polyurethanes have significant self-healing properties as well as high temperature endurance behavior around 1000°C. When exposed to sunlight, Xu et al.[34] developed crosslinked polyurethanes with a disulfide reactive group in the main chain that displayed self-healing and re-structuring characteristics. These reactive groups in the polymer chain (C—S or S—S) are easy to make and cause self-healing when exposed to light or heat. Zhao and colleagues[35] employed a photo-initiated thiol-epoxy addition technique in the presence of UV irradiation to synthesize photo-healable polymers from epoxidized oil in a single step. The basic raw material is epoxidized oil, and the photo-healing reactant is 7-mercapto-4-methyl coumarin, which grafts onto epoxidized triglycerides via its SH-groups. The thiol-epoxy addition mechanism and dimerization of coumarin reactive groups should yield photo-crosslinked and photo-healable polymers under UV light. It is concluded that following irradiation at 255 nm, the hardness of the polymer film decreases due to photocleavage of the grafted coumarin derivative into the graft monomer. After irradiation at $300 < \lambda < 450$ nm, the film's hardness may be restored, confirming the photo-reversible process.

3.2 | Microcapsule-based self-healing techniques

This kind of self-healing is achieved through two ways. In the first approach, the autonomous self-healing polymers have tiny microcapsules that hold the healing agent, which is unleashed into the polymer when the capsule coating is damaged. This procedure is known as the microcapsule-based self-healing mechanism. External stimuli such as light, pH, redox reaction, thermal heating, additives, catalyst, or solvents are used to initiate the healing response, which includes bond break and bond repair[36,37]

Second, by using a microcapsule-encapsulated self-healing technique, where an epoxy-based healing polymers were synthesized using cyclopentadiene like a healing agent and Grubb catalyst.[38] Since the connection between the liquid healing agent and the crosslinked thermoset is useless, and this healing technique is only acceptable for thermosetting polymers like epoxy, phenol formaldehyde (PF), and melamine (1,3,5-triamino-2,4,6-triazine) formaldehyde (MF). The main drawback of the capsule-based technique is that it can only treat the polymer surface once, whereas reversible breaking and repair of chemical

bonds-based healing approaches can provide a recyclable healing reaction.

3.3 | Polymer hydrogels for crack healing

Getachew and group designed hydrogels on the surface of porous structured poly(ether sulfonate) (PES) by using in-situ chemical graft polymerization of 2-acrylamido-2-methyl-1-propanesulfonic acid (AMPS).[39] In this mechanism, two interpretations have been shown. First, the presence of strong covalent bond between the hydrogel and substrates and second, the hydrogen bond rebuild after being disrupted. These results serve as a foundation for self-healing polymer.[40–43]

3.4 | Controlled/living polymerization-driven self-healing

In the polymerization process, chain-initiation is much quicker than chain propagation process. This approach has been used to make a variety of polymers with a narrow molecular weight distribution, and this unique characteristic has been used in a variety of applications.[44–47] Anionic, cationic, and radical polymerization are all followed by live polymerization. Living radical polymerization is demonstrated via reversible addition-fragmentation chain transfer (RAFT) and atom transfer radical polymerization (ATRP).[48–52] In the ATRP technique, the chain transfer and chain terminations are absent, and it contains of living chain end that again initiates the polymerization reaction with available reactive monomer.

Banerjee and team[53] fabricated the incorporating fluorescence-responsive ionic block copolymers (IBCPs) synthesized through a combination of ring-opening polymerization (ROP) of ε-caprolactone and xanthate-intermediated reversible addition-fragmentation chain transfer (RAFT) polymerization reactions yielded fluorescence-activated self-healing polymers. Polycaprolactone (PCL) was functionalized with xanthate to create a PCL-based macro-RAFT active agent, which was used to make block copolymers with cationic poly(2-[methcryloyloxy]ethyltrimethyl ammonium chloride) (PCL-b-PMTAC) and anionic poly(sodium-4-vinylbenzenesulfonate) (PCL-b-PMTAC) (PCL-b-PSS). They concluded that the healed polymer shows 550% elongation while 700% elongation shows uncut polymer.

3.5 | Schiff base reactions

To design the self-healing polymer, schiff base is known to be the best method and it is also used for the surface modification for many substrates.[54,55] Schiff base is a ligand that discovered by the condensation reaction of aldehyde with imines and used as a catalyst for many homogeneous and heterogeneous organic reactions. But Schiff base does not require any parameter due to its covalent type nature as it have its own efficiency to be healed by its own while other self-healing material requires some parameters like pH, light, etc.[56–60]

4 | APPLICATIONS OF SELF-HEALING MATERIALS

The auto healing characteristics of the self-healing materials when the structure encounter a drastic situation have made them suitable candidate for numerous applications in coatings, electronics, aerospace, medicals, construction, etc.

4.1 | Self-healing materials for space applications

In space environment spacecrafts suffer from harsh environmental variables which includes thermal cycles, UV radiation, space debris, mechanical vibrations, atomic oxygen, and many others.[61] In 1998, NASA Research announcement began the self-healing process. In starting, self-healing materials met with High technology readiness level, that is, TRL. Further, in 2001 NRA investigated the self-healing material by puncture test, fabricate the bladder, and verify the performance in a pressurized test. Finally, in 2005, research in self-healing has been sustained for the space applications. The requirements for inflatable structure are material property (such as low mass, thin, structural endurance, very low gas infusion, and flexibility etc.), operational lifetime (to preserve the material property for the maximum time and also to prevent the leakage), and the system level requirements (contain the need for compatible interfaces with the liner and restraint).[62] For example, a collision with tiny meteoroids and orbital debris which is also called micro meteoroid debris (MMOD) might produce punctures or cuts which causes depressurization and gas leakage.[62,63] Since no present system of space can be autonomously repaired, there is a disastrous failure. Structures, screens, damping elements and joints are frequently used in many spatial applications.[64] To recover or to handle these damages, materials have been used to survive these impact circumstances due to their lightweight, high mechanical characteristics, heat resistance and cost effectiveness.[65,66] In space, micrometeoroids and orbital debris are not the only risk to spacecraft integrity

and it is also affected by the other factors such as temperature fluctuations. To overcome this, Space self-healing composite are particularly intriguing, since they guarantee both autonomous repair and heat protection. Hence, they can be used in the production of circuits, rockets, and space mechanisms with high safety standards.[67]

In space application, there are three type of structural material are present which plays an important role in self-healing process. These materials are composite materials, ceramic material, and metal alloys for current conductor. The composite material, gives the broad use in self-healing application for coatings, actuators, compartments and antennas offering higher specific strength, fatigue resistance, improved corrosion resistance, and, also it is not applicable in application parts only, but also in some engine parts also. Composite materials used in space applications made up of polymer matrix reinforced with fiber like fiber-reinforced plastics (FRPs) reinforced with carbon fibers, fiber-reinforced plastics reinforced with glass fiber polymers (GFRPs), and carbon-epoxy fiber-reinforced plastic (CFRP) polymer composites such as thermosetting epoxy reinforced with high performance fibers.[68] Other composite materials include metal matrix composites, ceramic matrix composites (CMCs), and E-glass epoxy composites (EGCs).[66]

Ceramic materials are used for high temperature elements of thrusters and hard coating is also possible. And, metal alloys for current conductors includes wires, electronic circuitry. It is also based on intrinsic and extrinsic properties as well. However, composite material undergo structural damage due to the impact load. The damage gets initiated due to the formation of microscopic void which further leads to a deep microcracking and delamination. This leads to the premature failure of structure as represented in Figure 4a–d.[65]

From the last few researches, the cracks or damages was repaired by using the thermal plate, resin patches and injection methods. But these methods also have some disadvantages such as not suitable for invisible damages, needs damage monitors, and also not useful during the construction time. So, the use of material in aerospace application is limited because it requires more maintenance.

4.1.1 | Mechanical properties of materials and its effect on self-healing for space application

The mechanical properties of aerospace structural material are very important as it helps to determine the healing effectiveness of a self-healing materials. Mechanical properties used as a means of determining self-healing efficiency includes fracture toughness, impact strength, and fatigue property.

1. *Fracture property*: Fracture toughness is the ability of material with pre-cracks to resist fracture by absorbing energy. It is determined by fracture tests. Fracture toughness can be calculated by different types of mechanical tests such as compact tension, three-point bend, double-cantilever beam, tapered double-cantilever beam and width-tapered double-cantilever beam tests. The recovery rate of fracture toughness in self-healing material is around 30%–100%. According to White and team,[20] recovery rate should be around 75% of the fracture load by using the microcapsule containing epoxy resin. Many researches already have been taken place based on microcapsule healing technique. It was discovered that the amount of healing agent and catalyst used had a significant impact on fracture toughness recovery. According to Park et al.[70] and Chen et al.[71] demonstrated that the CFRP composites and 2MEP4F (a Diels–Alder crosslinked polymer formed from a monomer with four furan groups and another monomer with two maleimide groups) polymer systems both benefitted from reversible Diels–alder reactions. These crosslink polymeric materials contain 80% recovery rate and healing efficiency is good.
2. *Fatigue property*: Fatigue is the lowering of strength or the failure of a material due to repetitive stress, which may be above or below the yield strength. In general, polymers are less commonly employed for fatigue-loading applications than metallic materials. Hence, cyclic mechanical loadings can cause fatigue fractures in structural composite materials. So, self-healing material is the solution to these cases because these self-healing materials have the ability to heal cracks, and increase the fatigue life. Hamilton et al.[72] looked on the repair of fatigue damage in vascular epoxy composites. The material was subjected to extremely high stress levels, up to 80% of its quasistatic fracture toughness.
3. *Impact property*: Impacts property means the ability of the material to absorb energy during plastic deformation. Impact stresses can damage composite panels, which can be difficult to detect, posing a threat to aerospace composite structures. To address these challenges, heavy maintenance processes and nondestructive control methods have been devised, however the adoption of self-healing composite materials might be a simple and effective solution. Damage recovery after impact loading and impact properties have been the topic of a lot of research. Under low viscosity impact, Zainuddin and team[73] investigated the efficiency of healing agent under E-glass epoxy composites. It was

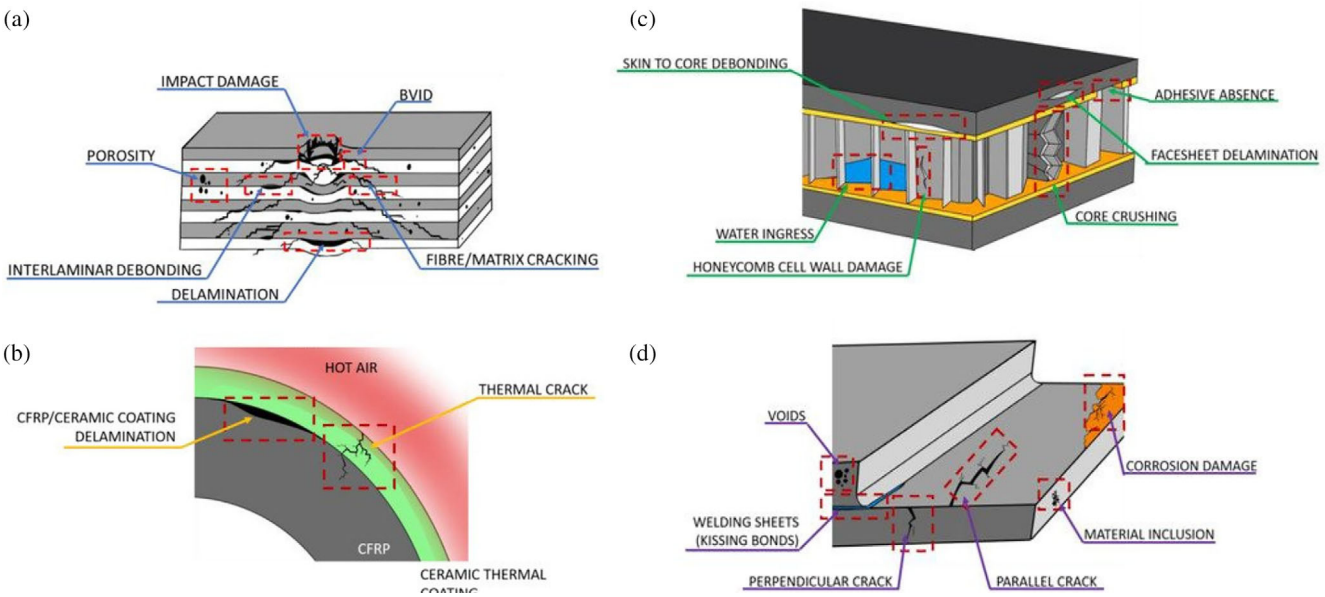


FIGURE 4 (a) Jet engine turbine blades, (b) honeycomb panels, (c) metallic aircraft, and (d) spacecraft components.[69] [Color figure can be viewed at wileyonlinelibrary.com]

shown that there is 53% improvement in peak load and around 86% improvement in energy to peak load. Under high-velocity impacts, Francesconi et al.[74] study shows that the replacement of ionomeric polymer to the aluminum panels for the protection of space debris and also shows that the aluminum behaves as the good fragmentation behavior.

4.1.2 | Self-healing polymers for space applications

In space industry, Self-healing polymers received much attention as they can lead to repair the structure autonomously after destructive damaged occurred by collisions with tiny meteoroids and orbital debris. This would improve security and permit long-term human expeditions to Mars and the Moon.[64] Xiang and group[75] designed the hybrid carbon fiber/epoxy composite with self-healing core shell nanofibers at interfaces by using co-electrospinning method. They concluded that formation of a core shell i.e. dicyclopentadiene/polyacrylonitrile (DCPD/PAN) by using dicyclopentadiene (DCPD) as a healing agent was encapsulated into polyacrylonitrile (PAN). This technique shows that the lightweight polymer matrix can be used in aerospace applications offering high-strength and high-toughness. It decreases the damages and regain the functionality and cycle life of the damage part. Polymer matrix composite found more reliable in aerospace because of its easy manufacturing and low cost. In general, self-healing applications in space industries also have to be constantly active and exposed to extreme environmental conditions due to certain difficulties.

In addition to polymers and composites, with the electropulsing treatment, metals and alloys might be exploited for the automatic localization of these materials for autonomous repairs in structural and electronic application. Like, in the ambient circumstances, Yu et al.[76] utilized electropulsing with the topmost current density of 4.1 kA/mm^2 to heal fractures in stainless steel SUS304, obtaining continuous healing areas on each side after seven cycles for discharge. These regions were considered by rapid melting and solidification and the formation of areas for fine recrystallation. Song et al.[77] utilized electropulsation to start titanium alloy healing.

Recent research by Laura et al. introduced the characterization on the intrinsic self-healing supramolecular polymer that is applicable for space applications also. In this study, they used a supramolecular polymer (REVERLINK® HR, ARKEMA) which has the strong irreversible covalent crosslinks and weak reversible chemical bond.[64] There has been an increasing research efforts in the space industry to fix the damages in space field. Composites made up of vinyl ester, 4,4′-Bismaleimidodiphenylmethane, dimethicone, Cyclopenta-1,3-diene, and CYTESTER (Cynate ester), and raw polymer like E-glass FRC and carbon FRC have been investigated.[66] Carbon fiber reinforced plastic is a self-healing polymer which were used in aeronautics by applying microcapsules which is loaded with healing materials for the purpose of preventing delamination fracture of Carbon fiber

reinforced plastic composites. The healing agent Cyclopenta-1,3-diene mixed up with polyepoxides by 20 wt.%. Another method also introduced to recover from exfoliate damages by using a thermoplastic polymer matrix, where Diels–alder reaction combined with ethylurea to accomplished the exfoliation inside the CFC (Carbon fiber composite) with the 75% and 85% healing efficiency.[78]

Carbon fiber reinforced plastic research has also been published on autonomic self-healing. The manufacture of fibers of 30–100 μm diameters, with an efficient volume of around 50%, was made by means of a customized hollow glass fiber (HGF). Either carbon fiber reinforced plastic (CFRP) and GFRP were implanted, and they were infused with uncured resin to provide the laminate with a self-healing feature. Some fibers packed with resin break away when damage is caused, the stored healing agent is released into a damaged site and the healing capabilities are recovered, that is, 89% of the laminate performance and 97% of the undamaged state were obtained.[79] The behavior of self-healing is generated by synthesis and encapsulation of Si $[OSn(n\text{-}C_4H_9)_2OOCCH_3]_4$, a highly effect catalyst for the curing of Polydimethylsiloxane.[80] Furthermore, Zhu et al. in 2019 introduced the UV- responsive self-healing microcapsules for aerospace coating. He reported that the UV-responsive microcapsule based coating for in-orbit damage repair was developed.[61]

As a result, coating becomes the important method in space industry as it involves fuselages, wing, engine as a defensive barrier. Brief discussion of self-healing polymers used in spacecraft given below in Table 4.

4.1.3 | Self-healing mechanism in space applications

For space application, there are two basic concepts which are existing on the mechanism of working of self-healing polymers are active and passive mechanism. Active mechanism should be ideal while passive mechanism (Health monitoring system) enter and trigger the self-healing and its power is available for activation. Passive mechanism is automatic and there is no requirement of infrastructure. Passive self-healing mechanism involves chemical reaction, environmental reaction, mechanical coverage, flowing materials and combined mechanism. Figure 5 represent the mechanism for space applications.[62]

Brief discussion on Passive self-healing mechanism[62,64]

1. *Chemical reactions*: The chemical reactions involve two materials which are fixed into a thin flexible film sheet. Upon penetration, these mixtures should mix and hence, void is sealed. As, a result this reaction should be an exothermic reaction. Reaction should be lightweight and not very thick, but also bears the hazards of limiting the shelf lives and the severity of the exothermic reaction.[62]
2. *Environmental reaction*: This reaction is comparable to the chemical reaction. But the difference is that it requires a single reactant which can be encapsulated and activated by the catalyst such as water and oxygen.[62]
3. *Mechanical coverage*: The concept of mechanical covering refers to the presence of material between the blade and the restraint of an inflated system on the bladder's pressure side. Mechanical coverage is low cost and lightweight technology.[62]
4. *Flowing materials*: This is one of the most-promising mechanism. Encapsulated materials such as viscoelastic gels, flows into the gap created in the barrier and develop the seal.[62]
5. *Combined mechanism*: It involves the basic concepts as we discussed above. In a hybrid concept, microvascular system is used. For flexible structure, microencapsulation is the technology that can be preferred. Like phosphorus, which is an environmental healing agent should be deposited in the microspheres insert in an independent film layer. Breakage of the bladder open the microsphere i.e. phosphorus to oxygen/water which causes the foaming reaction. Chemical reaction also exists in the microspheres.[62]

4.2 | Self-healing materials for construction applications

Concrete is the backbone of the construction field and the use of concrete has been taken from the ancient times.[71] Concrete is always being the first preference to construct the buildings, because of its exclusive nature like favorable fire resistance, high compressive strength, ease of casting, and cost efficiency. Concrete follows the positive properties, but a major worry is its low tensile strength which makes it susceptible to breaking. Damage and cracks in concrete not only comes from design flaws, and inappropriate selection of building materials, but also from environmental impacts like alkali-silica reaction.[82] Concrete cracking allows aggressive elements such as chloride, carbonate, or sulfate to penetrate the concrete. The progress of self-healing application in concrete depends on the immediately disruptive nature of the embedded microcapsules, the internal healing agent flow to the crack areas, and reaction occurs with catalyst.[83] The technology of self-healing in constructing the road is the better solution to encourage the strength and life of existing and future infrastructure. Asphalt binder is a

TABLE 4 Self-healing polymers in space application

Self-healing polymer	Mechanism	Method	Application	Reference
REVERLINK® HR, ARKEMA	Crosslinks irreversible and weak chemical reversible bond	Intrinsic	Supramolecular polymer	64
Cyclopenta-1,3-diene encapsulated microcapsule mixed with polyepoxides by 20 wt.%	Nil	Microcapsule	CFRP	78
Ethylurea	Diels–Alder reaction	Intrinsic	CFC	78
PDMS, adhesion promoter, dimethyldineodecanoate tin encapsulate in ethyl urea	$Si[OSn(n\text{-}C_4H_9)_2OOOCCH_3]_4$ synthesis that catalyzes polydimethylsiloxane treatment	Microcapsule	Nil	65
Epoxy silicon with titanium dioxide nanoparticles	Polymerization initiated by UV when UV light is exposed	Microcapsule	Coating	64
2,4,6-Tris([dimethylamino]methyl) phenol with Benzylidene-bis (tricyclohexylphosphino)-dichlororuthenium	ROMP	Microcapsule	Nil	81
Bicyclopentadiene and IPDI (isophorene diisocyanate)	Encapsulation in polymer fibers of self-healing materials	Encapsulation	FRPC	68
Core shell nanofiber of polyvinyl cyanide via co-axial electrospinning	In polyacrylonitrile were encapsulated two parts of the self-curing epoxy system with low viscosity epoxy resin	Encapsulation	PMC	80

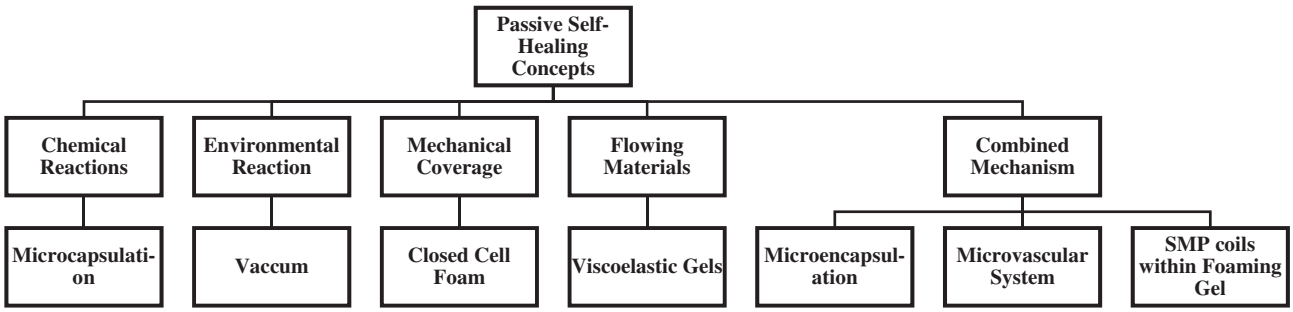


FIGURE 5 Taxonomy of passive self-healing concepts[62]

petroleum system product which generates petrol, diesel and many other fuels. Because of the viscoelastic property, its self-healing property may help to recover its originality by repairing the micro-crack providing the high durability.

Now a days, many researchers have been interested to discover the varieties of applications in microencapsulation of certain elements. The use of self-healing material in the concrete buildings might offer an alternate solution to concrete cracking and damages. Zhang and group[84] studied the self-healing capacity of the concrete specimens by using the non-ureolytic bacteria in expanded perlite (EP). They reported that as the time increases, the cracks in the concrete specimens with bacteria decreases and it happens when the carriers were immobilized. This study also concluded that when process evaluation is combined with microstructure analysis, EP particles may be used as innovative bacterium carriers in the healing of concrete fractures.

4.2.1 | Mechanical properties of materials and its effect on self-healing for construction application

1. *Fatigue life*: In the construction application, fatigue life is one of the main problems that affect the loading pattern in pavement. The relationship between the self-healing polymer and fatigue life of asphalt materials helps to prevent the maintenance and improve the asphalt pavement. The defects in the construction require a large amount of time and energy to restore

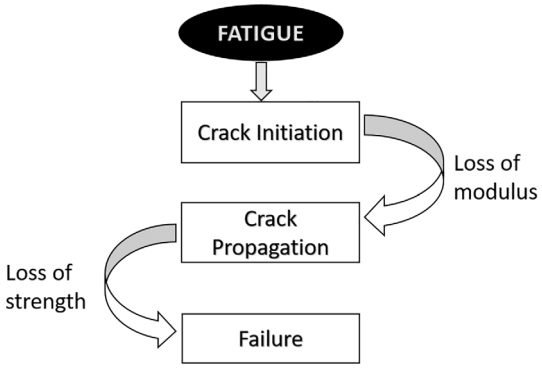


FIGURE 6 Fatigue three-step healing mechanism in construction of roads[85]

its original performance. Figure 6[85] explains the mechanism of fatigue life in the construction application. According to Tabaković et al.[86] the self-repairing property of a material is referred to as "its ability to substantially return to an original state or condition before exposure to a dynamic and susceptible environment by making the necessary adjustments to restore to normality and/or its ability to resist the formation of irregularities and/or defects. They concluded that 4-point bend fatigue test is the most suitable test for evaluating the performance of self-healing asphalt pavements.

4.2.2 | Self-healing polymers for construction application

In construction and building materials, self-healing polymers gives the best alternative to enhance the behavior of binders, binary cement (calcium sulfoaluminate [CSA] and Portland cement)[74] Many research theories have been taken to contact the adeptness of many self-healing polymers to adjust the defects in construction site. LV and team[87] shows the healing agents like poly(phenol–formaldehyde) (PF) resin shows that the PF material have the capacity to activate the resins from the crack growth. By using the situ polymerization process, it was employed to create PF microcapsules with a larger volume. It has been noticed that the size of microcapsules become larger and the thickness of the shell gets shrink. Smaller manufactured PF microcapsules have a higher possibility of being mechanically activated by the fracture than larger microcapsules.

Research also reported on the water permeability and rapid self-healing of sustainable sulfur composites by using superabsorbent polymer and binary cement by Seongwoo et al. The mechanism was detected by XRD method. The time taken for superabsorbent polymer was sealed by after 15 min. As a result, water penetration through cracks in sulfur composites would be successfully blocked off by applying appropriate portions of SAP and CSA agent.[88]

Researcher Sharareh et al. in 2019,[89] choose two binders on the basis of its viscoelastic nature, that is, PG 67-22 and PG 70-22 M with or without recycled asphalt materials, and 5% OXE-CHI-PUR (Oxetane-substituted Chitosan-Polyurethane) which is a self-healing material. Recycled asphalt shingles is the roofing material that uses asphalt for waterproofing while reclaimed asphalt pavements are used to removed or reprocessed pavement material having asphalt and aggregates. For the evaluation effect of self-healing material on the binder rheological properties, some tests also done which are the bending beam rheometer (BBR), multiple stress creep recovery (MSCR), dynamic shear rheometer (DSR), and by equating the superpave performance grade of the blends. By using the LAS test (Linear amplitude sweep), fatigue behavior of binder blends was shown.[89] Because of the traffic load, asphalt binder plays the crucial role. Self-healing polymer which is the new group polymer contains repeatable healing abilities. At molecular level, these repeatable healing process do not relate to any healing agent or catalyst but it requires different stimuli to produce self-healing polymer.[64] Further photochemical self-healing polymer is triggered by UV light and rebonding the active group has been triggered. For a UV light activated self-healing polymer, Polyurethane is a plastic material introduced by Ghosh et al.[90] By merging, polyurethane and OXE-CHI (oxetane-substituted chitosan) into a crosslinked polymer of OXE-CHI-PUR (Oxetane-chitosan-polyurethane). Micro-cracks have been appeared by containing UV light which leads to the polymer bond cleavage and free radicals. And these free radicals sealed the micro-cracks and hence, damage can be prepared.

In general, numerous self-healing polymer had been produced in the development of applications and are summed up in Table 5.

4.2.3 | Self-healing mechanism in construction applications

Microcapsules are produced by the physical process or mechanical action via physical techniques. To operate the usage of aquatic bacteria-based treatments to self-heal Marine Concrete subjected to extreme sea water was reported by Palin et al.[92] In 2017, A. Kanellopoulos searched that bacterial-based agents synthesized microcapsules of gelatin and acacia gum to contain sodium silicate liquid for self-healing concrete production found

TABLE 5 Self-healing materials in construction applications

Self-healing agent	Application	Findings	Reference
SAP and binary cement (CSA+ Portland cement)	Self-healing of sustainable sulfur composites	The presence of SAP chromatography significantly increased the fast self-selling performance	88
Non-ureolytic bacteria, that is, Bacillus cohnii, immobilized in EP	Self-healing method for cementitious materials	As the time increases, the cracks in the concrete specimens with bacteria decreases and it happens when the carriers were immobilized.	84
Encapsulation materials containing polyurethane	Self-healing method for cementitious materials	By performing water permeability tests, it has been shown that the efficiency of crack repairing has been improved and showed the improvement in water tightness	91

very balanced.[93] To put the self-healing microcapsules in concrete, microencapsules were prepared at first. The embedded self-healing microcapsules are susceptible to concrete cracking, it tends to break away and releases the encapsulated heals. The healing agent polymerizes and acts to seal and repair cracks when it is exposed to the integrated catalyst.[94]

The successful creation of the mechanism of self-healing depends on three factors[83]:

- breaking of embedded microcapsules
- releasing of healing agent into the cracks
- polymerization of self-healing agent with catalyst

5 | CONCLUSION AND FUTURE ASPECTS

The self-healing polymers are versatile and can be used as a self-healing agent for applications in space, biomedical, construction, automobiles etc. Literature studies have shown that these self-healing polymeric materials are promising materials as they help repair structural damages, cracks or defects in materials through different polymerization reactions with excellent mechanical, physical, rheological, thermal property which makes them more demanding for various applications. This article gives an overview of self-healing polymers used in space and construction applications and how they demonstrate considerable improvements on the crack properties of stressed structures in these applications. The mechanism involved during the self-healing processes is much favorable in the application because it increases the life cycle of the material. Although, some environmental factors such as temperature, pressure, light etc. do offer some challenges which has to be addressed carefully during the different phases of self-healing material in order to achieve the long-time service life. Thus, in future performance properties like cracking at low temperature, rutting, and self-healing ability should be evaluated. There is a requirement for more advanced healing techniques that extend beyond pure intrinsic and extrinsic healing and more emerging polymeric materials should also be explored. However, this study is still in progress and further innovations are needed. In future investigations, test criteria for this must be developed.

AUTHOR CONTRIBUTIONS

Megha Goyal: Writing – original draft (lead). **Nitu Bhatnagar:** Writing – review and editing (equal).

ACKNOWLEDGMENT

The authors gratefully acknowledge the facility provided by the Manipal University Jaipur.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Nitu Bhatnagar https://orcid.org/0000-0001-6250-8940

REFERENCES

[1] F. Bueche, W. Cashin, P. Debye, *J. Chem. Phys.* **1952**, *20*, 1956.
[2] V. Kargin, I. Malinskii, V. Prokopenko, *Doklady* **1969**, *189*, 568.
[3] Y. M. Malinskii, V. Prokopenko, N. Ivanova, V. Kargin, *Polym. Mech.* **1970**, *6*, 240.
[4] K. Jud, H. Kausch, J. Williams, *J. Mater. Sci.* **1981**, *16*, 204.
[5] R. Wool, K. O'Connor, *J. Polym. Sci., Polym. Lett. Ed.* **1982**, *20*, 7.
[6] S. Wang, M. W. Urban, *Nat. Rev. Mater.* **2020**, *5*, 562.
[7] Y. Yang, D. Davydovich, C. C. Hornat, X. Liu, M. W. Urban, *Chem* **2018**, *4*, 1928.
[8] Y. Chen, A. M. Kushner, G. A. Williams, Z. Guan, *Nat. Chem.* **2012**, *4*, 467.

[9] J. Nji, G. Li, *Polymer* **2010**, *51*, 6021.
[10] C. C. Corten, M. W. Urban, *Adv. Mater.* **2009**, *21*, 5011.
[11] Y. Yang, M. W. Urban, *Angew. Chem., Int. Ed.* **2014**, *53*, 12142.
[12] H. Ying, Y. Zhang, J. Cheng, *Nat. Commun.* **2014**, *5*, 1.
[13] X. Chen, M. A. Dam, K. Ono, A. Mal, H. Shen, S. R. Nutt, K. Sheran, F. Wudl, *Science* **2002**, *295*, 1698.
[14] B. Ghosh, M. W. Urban, *Science* **2009**, *323*, 1458.
[15] K. Imato, M. Nishihara, T. Kanehara, Y. Amamoto, A. Takahara, H. Otsuka, *Am. Ethnol.* **2012**, *124*, 1164.
[16] P. Cordier, F. Tournilhac, C. Soulié-Ziakovic, L. Leibler, *Nature* **2008**, *451*, 977.
[17] M. Burnworth, L. Tang, J. R. Kumpfer, A. J. Duncan, F. L. Beyer, G. L. Fiore, S. J. Rowan, C. Weder, *Nature* **2011**, *472*, 334.
[18] M. Nakahata, Y. Takashima, H. Yamaguchi, A. Harada, *Nat. Commun.* **2011**, *2*, 1.
[19] M. W. Urban, D. Davydovich, Y. Yang, T. Demir, Y. Zhang, L. Casabianca, *Science* **2018**, *362*, 220.
[20] S. R. White, N. R. Sottos, P. H. Geubelle, J. S. Moore, M. R. Kessler, S. Sriram, E. N. Brown, S. Viswanathan, *Nature* **2001**, *409*, 794.
[21] M. Kessler, N. R. Sottos, S. White, *Compos. Part A Appl. Sci. Manuf.* **2003**, *34*, 743.
[22] P. C. Je, M. T. Sultan, C. P. Selvan, S. Irulappasamy, F. Mustapha, A. A. Basri, S. N. Safri, *J. Mater. Res. Technol.* **2020**, *9*, 7370.
[23] N. N. F. Nik Md Noordin Kahar, A. F. Osman, E. Alosime, N. Arsat, N. A. M. Azman, A. Syamsir, Z. Itam, Z. A. A. Hamid, *Polymer* **2021**, *13*, 1194.
[24] A. B. Brochu, S. L. Craig, W. M. Reichert, *J. Biomed. Mater. Res. A* **2011**, *96*, 492.
[25] D. Cadogan, C. Scheir, A. Dixit, J. Ware, E. Cooper, P. Kopf, 47th AIAA/ASME/ASCE/AHS/ASC Struct., Struct. Dynam., Mater. Conf. 14th AIAA/ASME/AHS Adapt. Struct. Conf. 7th **2006**, p. 1897.
[26] Q. Zhang, L. Liu, C. Pan, D. Li, *J. Mater. Sci.* **2018**, *53*, 27.
[27] S. Van Der Zwaag, Springer, Delft **2007**, *100*, 1.
[28] S. Thakur, *Sustainable Nanostructural Materials for Shape Memory, Self-healing, and Self-cleaning Applications*, Elsevier, Tezpur University, Tezpur, Assam **2019**, p. 235.
[29] S. Utrera-Barrios, R. Verdejo, M. A. López-Manchado, M. H. Santana, *Mater. Horiz.* **2020**, *7*, 2882.
[30] I. L. Hia, V. Vahedi, P. Pasbakhsh, *Polym. Rev.* **2016**, *56*, 225.
[31] C. Dry, *Compos. Struct.* **1996**, *35*, 263.
[32] K. R. Reddy, A. El-Zein, D. W. Airey, F. Alonso-Marroquin, P. Schubel, A. Manalo, *Nano-Struct. Nano-Objects* **2020**, *23*, 100500.
[33] C. Cheng, J. Li, F. Yang, Y. Li, Z. Hu, J. Wang, *J. Polym. Res.* **2018**, *25*, 1.
[34] W. M. Xu, M. Z. Rong, M. Q. Zhang, *J. Mater. Chem. A* **2016**, *4*, 10683.
[35] Y. H. Zhao, D. Vuluga, L. Lecamp, F. Burel, *RSC Adv.* **2016**, *6*, 32098.
[36] Q. Yan, A. Feng, H. Zhang, Y. Yin, J. Yuan, *Polym. Chem.* **2013**, *4*, 1216.
[37] L.-P. Lv, Y. Zhao, N. Vilbrandt, M. Gallei, A. Vimalanandan, M. Rohwerder, K. Landfester, D. Crespy, *J. Am. Chem. Soc.* **2013**, *135*, 14198.
[38] K. S. Toohey, C. J. Hansen, J. A. Lewis, S. R. White, N. R. Sottos, *Adv. Funct. Mater.* **2009**, *19*, 1399.
[39] B. A. Getachew, S.-R. Kim, J.-H. Kim, *Environ. Sci. Technol.* **2017**, *51*, 905.
[40] H. Zhang, H. Xia, Y. Zhao, *ACS Macro Lett.* **2012**, *1*, 1233.
[41] J. Cui, A. del Campo, *Chem. Commun.* **2012**, *48*, 9302.
[42] D. C. Tuncaboylu, A. Argun, M. P. Algi, O. Okay, *Polymer* **2013**, *54*, 6381.
[43] B. K. Ahn, D. W. Lee, J. N. Israelachvili, J. H. Waite, *Nat. Mater.* **2014**, *13*, 867.
[44] Z. Zhao, H. Shen, K. Sui, G. Wang, *Polymer* **2018**, *137*, 364.
[45] J. Huang, Z. Liu, D. Cui, X. Liu, *ChemCatChem* **2018**, *10*, 42.
[46] S. Li, D. Liu, Z. Wang, D. Cui, *ACS Catal.* **2018**, *8*, 6086.
[47] W. Fan, M. Tosaka, S. Yamago, M. F. Cunningham, *Angew. Chem., Int. Ed.* **2018**, *57*, 962.
[48] J. Chen, M. Liu, Q. Huang, L. Huang, H. Huang, F. Deng, Y. Wen, J. Tian, X. Zhang, Y. Wei, *Chem. Eng. J.* **2018**, *337*, 82.
[49] Q. Wan, Q. Huang, M. Liu, D. Xu, H. Huang, X. Zhang, Y. Wei, *Appl. Mater. Today* **2017**, *9*, 145.
[50] R. Jiang, H. Liu, M. Liu, J. Tian, Q. Huang, H. Huang, Y. Wen, Q.-Y. Cao, X. Zhang, Y. Wei, *Mater. Sci. Eng. C* **2017**, *81*, 416.
[51] X. Zhang, K. Wang, M. Liu, X. Zhang, L. Tao, Y. Chen, Y. Wei, *Nanoscale* **2015**, *7*, 11486.
[52] Y. Liu, L. Mao, X. Liu, M. Liu, D. Xu, R. Jiang, F. Deng, Y. Li, X. Zhang, Y. Wei, *Mater. Sci. Eng. C* **2017**, *79*, 590.
[53] S. L. Banerjee, R. Hoskins, T. Swift, S. Rimmer, N. K. Singha, *Polym. Chem.* **2018**, *9*, 1190.
[54] S.-H. Lee, S.-R. Shin, D.-S. Lee, *Mater. Des.* **2019**, *172*, 107774.
[55] C. Luo, M. Li, R. Yuan, Y. Yang, Z. Lu, L. Ge, A. C. S. Appl, *Bio Mater.* **2020**, *3*, 1481.
[56] M. Vahedi, J. Barzin, F. Shokrolahi, P. Shokrollahi, *Macromol. Mater. Eng.* **2018**, *303*, 1800200.
[57] J. Ren, H. Xuan, L. Ge, *Appl. Surf. Sci.* **2019**, *469*, 213.
[58] L. Tang, X. Chen, L. Wang, J. Qu, *Polym. Chem.* **2017**, *8*, 4680.
[59] J. Ren, H. Xuan, L. Ge, *Eur. Polym. J.* **2017**, *93*, 521.
[60] H. Li, J. Bai, Z. Shi, J. Yin, *Polymer* **2016**, *85*, 106.
[61] Y. Zhu, K. Cao, M. Chen, L. Wu, A. C. S. Appl, *Mater. Interfaces* **2019**, *11*, 33314.
[62] J. Ferl, J. Ware, D. Cadogan, J. Yavorsky, *SAE Int.* **2007**, *1*, 3211. https://doi.org/10.4271/2007-01-3211.
[63] E. Buslov, I. Komarov, V. Selivanov, V. Titov, N. Tovarnova, V. Feldstein, *Acta Astronaut.* **2019**, *163*, 54.
[64] L. Pernigoni, A. M. Grande, *Acta Astronaut.* **2020**, *177*, 697.
[65] M. W. Lee, *Polymer* **2020**, *12*, 379.
[66] R. Das, C. Melchior, K. Karumbaiah, *Self-healing Composites for Aerospace Applications*, Elsevier, University of Auckland, Auckland, New Zealand **2016**, p. 333.
[67] L. Pernigoni, U. Lafont, A. M. Grande, *CEAS Space J.* **2021**, *13*, 341.
[68] S. Sinha-Ray, D. Pelot, Z. Zhou, A. Rahman, X.-F. Wu, A. Yarin, *J. Mater. Chem.* **2012**, *22*, 9138.
[69] F. Ciampa, P. Mahmoodi, F. Pinto, M. Meo, *Sensors* **2018**, *18*, 609.
[70] J. S. Park, H. S. Kim, H. T. Hahn, *Compos. Sci. Technol.* **2009**, *69*, 1082.
[71] X. Chen, F. Wudl, A. K. Mal, H. Shen, S. R. Nutt, *Macromolecules* **2003**, *36*, 1802.
[72] A. Hamilton, N. R. Sottos, S. R. White, *Polymer* **2012**, *53*, 5575.

[73] S. Zainuddin, T. Arefin, A. Fahim, M. Hosur, J. Tyson, A. Kumar, J. Trovillion, S. Jeelani, *Compos. Struct.* **2014**, *108*, 277.
[74] A. Francesconi, C. Giacomuzzo, A. M. Grande, T. Mudric, M. Zaccariotto, E. Etemadi, L. Di Landro, U. Galvanetto, *Adv. Space Res.* **2013**, *51*, 930.
[75] X. F. Wu, A. Rahman, Z. Zhou, D. D. Pelot, S. Sinha-Ray, B. Chen, S. Payne, A. L. Yarin, *J. Appl. Polym. Sci.* **2013**, *129*, 1383.
[76] T. Yu, D. Deng, G. Wang, H. Zhang, *J. Cleaner Prod.* **2016**, *113*, 989.
[77] H. Song, Z.-J. Wang, X.-D. He, J. Duan, *Sci. Rep.* **2017**, *7*, 1.
[78] G. Williams, R. Trask, I. Bond, *Compos. A Appl. Sci. Manuf.* **2007**, *38*, 1525.
[79] I. Bond, G. Williams, R. Trask, 22nd Tech. Conf. Am. Soc. Compos. 2007-Compos.: Enabl. N. Era Civil Aviat. **2007**, pp. 1960–1974.
[80] R. E. Neisiany, J. K. Y. Lee, S. N. Khorasani, S. Ramakrishna, *J. Appl. Polym. Sci.* **2017**, *134*, 44956.
[81] L. Guadagno, M. Raimondo, G. Iannuzzo, S. Russo, *AIP Conf. Proc.* **2010**, *1255*, 267.
[82] Z. Itam, S. Beddu, D. Mohammad, N. L. M. Kamal, N. Razak, Z. Hamid, *Mater. Today: Proc.* **2019**, *17*, 717.
[83] F. Gilabert, K. Van Tittelboom, E. Tsangouri, D. Van Hemelrijck, N. De Belie, W. Van Paepegem, *Cem. Concr. Compos.* **2017**, *79*, 76.
[84] J. Zhang, Y. Liu, T. Feng, M. Zhou, L. Zhao, A. Zhou, Z. Li, *Constr. Build. Mater.* **2017**, *148*, 610.
[85] D. Sun, G. Sun, X. Zhu, A. Guarin, B. Li, Z. Dai, J. Ling, *Adv. Colloid Interface Sci.* **2018**, *256*, 65.
[86] A. Tabaković, L. Schuyffel, A. Karač, E. Schlangen, *Appl. Sci.* **2017**, *7*, 647.
[87] L. Lv, E. Schlangen, Z. Yang, F. Xing, *Materials* **2016**, *9*, 1025.
[88] S. Gwon, E. Ahn, M. Shin, *Constr. Build. Mater.* **2020**, *265*, 120306.
[89] S. Shirzad, M. M. Hassan, M. A. Aguirre, L. N. Mohammad, S. Cooper Jr., I. I. Negulescu, *Constr. Build. Mater.* **2019**, *220*, 187.
[90] B. Ghosh, K. V. Chellappan, M. W. Urban, *J. Mater. Chem.* **2011**, *21*, 14473.
[91] K. Van Tittelboom, E. Tsangouri, D. Van Hemelrijck, N. De Belie, *Cem. Concr. Compos.* **2015**, *57*, 142.
[92] D. Palin, V. Wiktor, H. Jonkers, *Smart Mater. Struct.* **2016**, *25*, 084008.
[93] A. Kanellopoulos, P. Giannaros, D. Palmer, A. Kerr, A. Al-Tabbaa, *Smart Mater. Struct.* **2017**, *26*, 045025.
[94] C. Xue, W. Li, J. Li, V. W. Tam, G. Ye, *Struct. Concr.* **2019**, *20*, 198.

AUTHOR BIOGRAPHIES

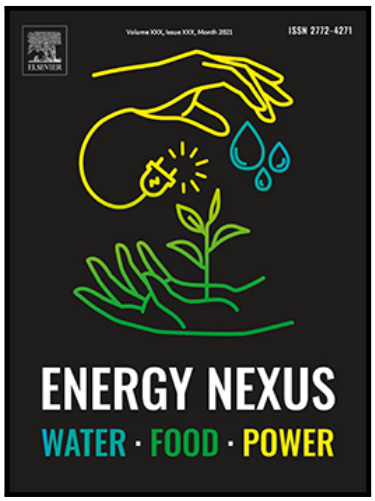
Ms. Megha Goyal did her Master's in inorganic chemistry from the department of Chemistry, JECRC University, Jaipur, Rajasthan in 2018. Presently, she is pursuing her Ph.D. from Manipal University Jaipur, Rajasthan under the supervision of Dr. Nitu Bhatnagar. Her research interests include the coating of conductive polymers on Li-ion battery.

Mr. Satya Narayan Agarwal is Assistant Professor at Department of Electrical Engineering, School of Electrical, Electronics & Communication Engineering, Manipal University Jaipur. He received his Master of Technology in Instrumentation & Control Engineering from Delhi Technological University, Delhi. He received his Bachelor of Technology in Electronics Instrumentation & Control Engineering, University College of Engineering, Rajasthan Technical University Kota. His research area is energy storage devices and its application, control systems and instrumentation.

Dr Nitu Bhatnagar is currently working as Registrar of Manipal University Jaipur, with effect from August 01, 2021. She basically belongs to Department of Chemistry, School of Basic Sciences, MUJ. She did her BSc and MSc from Pt. RSS University, Chhattisgarh and PhD from Sikkim Manipal Institute of Technology, SMU, Sikkim. During her PhD, she received the prestigious Women Scientist Fellowship from Department of Science & Technology, Govt. of India. She has more than 15 years of teaching and research experience. She has published 5 book chapters and 28 research articles in international journals of repute, including 16 SCI/Scopus publications. She has delivered many lectures as keynote/Invited speaker. She has presented more than 30 research papers in various conferences in India & Abroad. She has been Invited as Visiting Faculty at CUHK, Hong Kong under Academic Visitors Scheme of CUHK. She has collaborated with many researchers in India and abroad. Her area of research includes nanocomposites, ayurvedic bhasma, plasma surface modification and adhesive technology. One student has been awarded PhD Degree under her guidance; one has submitted his thesis, & three research scholars are pursuing their PhD degree. She has a copyright granted on "Academic Audit Framework".

How to cite this article: M. Goyal, S. N. Agarwal, N. Bhatnagar, *J. Appl. Polym. Sci.* **2022**, e52816. https://doi.org/10.1002/app.52816

Journal Pre-proof



Algal Bioremediation of heavy metals: an insight into removal mechanisms, recovery of by-products, challenges, and future opportunities

Mohita Chugh , Lakhan Kumar ,
Dr Maulin P Shah Chief Scientist & Head ,
Dr Navneeta Bharadvaja

PII: S2772-4271(22)00084-5
DOI: https://doi.org/10.1016/j.nexus.2022.100129
Reference: NEXUS 100129

To appear in: *Energy Nexus*

Received date: 24 March 2022
Revised date: 16 July 2022
Accepted date: 7 August 2022

Please cite this article as: Mohita Chugh , Lakhan Kumar , Dr Maulin P Shah Chief Scientist & Head , Dr Navneeta Bharadvaja , Algal Bioremediation of heavy metals: an insight into removal mechanisms, recovery of by-products, challenges, and future opportunities, *Energy Nexus* (2022), doi: https://doi.org/10.1016/j.nexus.2022.100129

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license
(http://creativecommons.org/licenses/by-nc-nd/4.0/)

Journal Pre-proof

Highlights

- Algae can be used as a phycoremediator of heavy metals and they show high resistance towards them.
- Microalgae show various Self-defense mechanisms against heavy metal toxicity.
- To enhance sustainability of phycoremediation, different integrated systems can be used for the remediation of heavy metals.

Journal Pre-proof

Journal Pre-proof

Algal Bioremediation of heavy metals: an insight into removal mechanisms, recovery of by-products, challenges, and future opportunities

Mohita Chugh[1], Lakhan Kumar[1], Maulin P Shah[2], Navneeta Bharadvaja[1]*

Author 1: Mohita Chugh, Department of Biotechnology, Delhi Technological University, Delhi-110042, India.

ORCID ID: https://orcid.org/0000-0002-0175-0546

E: mohita.2912@gmail.com

Author 2: Lakhan Kumar, Department of Biotechnology, Delhi Technological University, Delhi-110042, India.

ORCID ID: https://orcid.org/0000-0002-4971-3023

E: adarsh.lakhan@gmail.com

Author 3: Dr Maulin P Shah, Chief Scientist & Head, Industrial Wastewater Research Lab, India

ORCID ID: https://orcid.org/0000-0003-1880-5221

E: maulinshah1979@gmail.com; hahmp@beil.co.in

Author 4: Dr Navneeta Bharadvaja, Department of Biotechnology, Delhi Technological University, Delhi-110042, India.

ORCID ID: https://orcid.org/0000-0001-6727-9622

E: navneetab@dce.ac.in

***Corresponding Author Email:** navneetab@dce.ac.in

Journal Pre-proof

Journal Pre-proof

Journal Pre-proof

Abstract

Recently, modern techniques have gained attention for remediation of heavy metals present in the wastewater. Biological processes have been explored and employed as they have been proven to be sustainable, simple, and green. With the help of this review, we have tried to bring out knowledge on different microalgae strains used for the removal of heavy metals and their mechanism of action for the remediation. The removal of heavy metals is performed by microalgal species through different extracellular and intracellular mechanisms. Phycoremediation of heavy metal mechanisms has been discussed in detail. It is seen that alone bioremediation of heavy metals is not an economical strategy. Thus, efforts on the development of integrated systems have been increasing to make this process more economical and sustainable. This review further elaborates upon recent advancements in the remediation of heavy metals using microalgae. Along with this, challenges to phycoremediation of heavy metals and opportunities to overcome have been discussed which can make this process more efficient, economical, sustainable, and clean technology. Through the discussions in this review, it can be concluded that phycoremediation can play a significant role in remediation of heavy metals in a sustainable way along with realization of bioeconomic potential.

Keywords

Phycoremediation, Different types of mechanisms, Algae tolerance for Heavy metal, Heavy Metals, Economical Integrated Set-up.

Table of Contents Page No.

1. Introduction 4
2. Heavy metals: sources and hazardous effects 5
3. Phycoremediation: an insight to the mechanism 8
   3.1 Self-defense mechanisms of algae against heavy metal toxicity 9
4. Phycoremediation of heavy metals: Cadmium, Arsenic, Chromium, and Arsenic 12
5. Phycoremediation: a sustainable remediation technique 17
   5.1 Phycoremediation of heavy metals along with production of Bioenergy 18

Journal Pre-proof

6. Recent advancements in phycoremediation approach towards heavy metals removal from wastewater 19
7. Challenges and Opportunities 21
8. Conclusion 23

Journal Pre-proof

Journal Pre-proof

1. Introduction

Heavy metals are considered to be one of the most toxic water pollutants. Still, 80% of the untreated wastewater is released in the water bodies including freshwater bodies which are used for domestic purposes. This is leading to global water stress because of the increasing scarcity of freshwater resources. A study estimates that approximately sixty percent of the global population will suffer from water stress by the year 2025 (Khalid et al., 2018). Heavy metals present in the wastewater arise from various naturally occurring processes (like volcanic eruptions or other) and anthropogenic activities (like industrial solid and liquid waste discharges). Heavy metals are released into the environment by several industries including electroplating, textiles, dyes, mining, batteries, automobiles, and many others (Kumar, Khushbu, et al., 2022). From these industries, heavy metals are released into the air, soil and finally into waterbodies (Kesari et al., 2021). Heavy metals like mercury, lead, arsenic, nickel, chromium, titanium, cadmium, molybdenum, copper, zinc, nickel, manganese, iron, cobalt, boron, silver, and gold etc., enter the food chain and get accumulated inside the human body and consequently cause adverse effects on the human health (Pacheco et al., 2020) (Billah et al., 2017) (UNESCO, 2015). Apart from human health, these pollutants also show adverse impacts on flora and fauna (Coelho et al., 2016). This environmental degradation raised the demand for devising remediation strategies.

Traditional and modern techniques are used for the removal of heavy metals present in the wastewater. Traditional techniques including ion exchange, electrochemical treatment, osmosis, evaporation, and precipitation require more energy input making these techniques more expensive. Further, most of these methods pose challenges towards sustainable removal of metals from the environment (Nguyen et al., 2020) (Priyadarshani et al., 2012). To overcome this, modern techniques like biological methods are used which are more economical and sustainable approach for the removal of heavy metals (Kumar, Chugh, et al., 2022). At the same time, these methods are easy to apply. The term phycoremediation refers to the remediation of water pollutants using algae. It is a potential technique for the abatement of heavy metals as algae are abundantly available, inexpensive, have excellent metal removal capacity, eco-friendly and give rise to value-added products. Algae are generally classified into two types i.e., macroalgae and microalgae. Macroalgae are of three types: brown, red and green whereas microalgae are majorly classified into four main types which are diatoms, green, golden and blue-green algae (Ali Redha, 2020). During phycoremediation, most of the pollutants are removed via biosorption, bioaccumulation, and detoxification

Journal Pre-proof

mechanism shown by algae. In biosorption, sorption of the heavy metals on the biomass surface occurs irrespective of the fact that it is living or not (Ugya et al., 2021). In bioaccumulation, heavy metals are transported inside the algal cell. In detoxification, various phyco-chelators released by algae helps in the transformation of toxic heavy metal into non-toxic compound (Kikuchi & Tanaka, 2012).

Form sustainability point of view, phycoremediation alone isn't a best choice. In recent times, algal biomass have been explored for the production of various economical products such as bio-oil, nanoparticles, bioenergy and bio-fertilizer (Kumar & Bharadvaja, 2021). Algal based bioenergy is observed to be the most eco-friendly and sustainable energy (Vassilev & Vassileva, 2016). Biomass obtained during phycoremediation can be used towards production of bioenergy and several other high value products to make it sustainable (Forján et al., 2015). It can lead to the production of green energy along with the remediation of heavy metals present in the wastewater. This review will highlight the current understanding of the phycoremediation of heavy metals along with the in-depth knowledge of major mechanism through which algae remediate heavy metals from wastewater. Descriptive information of the sources of heavy metals along with their hazardous effects is also given. Furthermore, a brief discussion on effect of heavy metals on the productivity of commercial product and how to enhance the accumulation of these by-products along with the remediation of heavy metals is also presented. To make this technique more efficient; economical integrated systems, recent advancements, challenges and strategies to carry out in the future are provided and explored.

2. Heavy metals: sources and hazardous effects

Globally, heavy metals are one of the most hazardous contaminants responsible for environmental damage. Industries like batteries, automobiles, mining, smelting of ores, agricultural waste like fertilizers and pesticides release heavy metals like arsenic, lead, chromium, cadmium, mercury, nickel, silver, and uranium, etc., into the environment. Recently, barium is also added to the list. (Usmani et al., 2020)(Jie et al., 2021). These are non-biodegradable in nature and have a hazardous impact on the whole ecosystem. Heavy metals are indiscriminately used for anthropogenic purposes which alter the bio-geochemical cycles. The presence of the non-permissible amount of heavy metals in the soil, air and water causes harmful effects on the ecological diversity of the earth (Nájera-Martínez et al., 2021)(Yasin et al., 2021). Their accumulation into the environment leads to biomagnification (Ali et al., 2019). Heavy metals also tends to cause respiratory, renal, mental and cardiovascular

Journal Pre-proof

disorders and leads to cancer (Chugh et al., 2022). Here in this section, we have discused four most lethal heavy metals arsenic, chromium, cadmium, and mercurry.

2.1 Arsenic

Arsenic is considered to be highly toxic heavy metal as at a lower concentraion in the range of 0.01 µg/mL, it is hazardous and is considered to be carcinogenic. Major routes from where arsenic enters the human body is through inhalation, skin, and ingestion. Then it affects lung, heart, kidney, bladder, skin, and the central nervous system. It causes arsenicosis, leucomelanosis, hyperkerotosis, arsenical dermatitis, diabetes, abortions, and alter the immune system. It impacts the cellular metabolism of plants. It imbalances the phosphate in DNA and ends up disturbing the balance of carbon, amino acid, proteins, nitrogen, sulphur, and other metabolic pathways (Finnegan & Chen, 2012). Arsenic cause decrease in its growth, photosynthetic activity as well as disrupts thylakoid membrane (Atul Kumar Upadhyay et al., 2020). In microbes it destroy cell organelles, distrupts protein, DNA, lipid, and proteins (A. K. Upadhyay et al., 2016). It is released through anthropogenic activities like mining, smelting, medical, chemical and agricultural waste. It is present in fungicide, pesticides, and herbicides. It is used in the alloy manufacturing processes (Arora et al., 2017). Two forms of arsenic i.e. arsenate and arsenite which are easily uptake by plants and eventually enters the food chain.

2.2 Cadmium

It is one of the most invasive heavy metals. It is released into the environment through natural as well widespread use of industrial activities. Cadmium is naturally released through the mineral or rocks. Cadmium enters the food chain and then to the gastrointestinal tract, affecting organs like heart, kidney, liver, and reproductive organs. It is considered to be teratogenic and carcinogenic. Cadmium is inhaled by the humans and it affects the chest and lungs. It causes nausea, vomiting, cough and other complications. It affects the growth, morphology as well as cellular metabolism, photosynthesis in plants and microbes and sometimes leads to death (Xie et al., 2021) (Chandrashekharaiah et al., 2021). Cadmium is used as a stabilizer, or present in the dyes, fertilizers, batteries such as nickel-cadmium battery, fuel combustion and act as a corrosive agent (Genchi et al., 2020). It is a non-essential heavy metal for humans. Industries like metal industries, electroplating, smelting, plastic, dyeing and refining industries leads to the release of cadmium into the environment. Agricultural sector including fertilizer and pesticide

Journal Pre-proof

industries, semiconductor, welding, cigarette industries, and nuclear plants release cadmium into the environment (Kumar & Bharadvaja, 2020b).

2.3 Chromium

Chromium is present in multiple valence states in the environment. Chromium (VI) is more lethal than chromium (III) as it is carcinogenic in nature as compared to Chromium (III) which is immobile and poorly soluble. Chromium (VI) can easily enters the cell membrane and causes damage to genetic materials, skin, lungs, kidney, liver, and digestive system (Leong & Chang, 2020). It causes coughing, wheezing, ulcers, edema, haemorrhage, and chronic bronchitis (Elahi et al., 2020). In microbes it induces phycochelators against the toxicity of chromium (Aharchaou et al., 2017). The major origin of the chromium in the environment is through the natural process because of the chromite ore. It is also present in the sewage sludge, electroplating, and tannery effluent. When these contaminants along with the ore itself are not disposed of properly, it leads to the leaching of the contaminants into the surface water and groundwater (Tumolo et al., 2020). Industries like steel, dyes, paints, leather, electropainting, asbestos, cement, chemical industries are also responsible for the release of chromium into the environment.

2.4 Mercury

The widespread use of mercury in industries like mining, smelting, coals, electronics, pesticides, and chemical etc., is the main reason behind its release into the environment (Peng et al., 2017). Mercury is added in the agricultural land through the manures, sludge, fungicides or fertilizers. Mercury is converted into neurotoxin methyl mercury. It causes nervous disorders, gene expression alterations, autoimmune diseases, depression, gingivitis, Minamata, irritation in eyes, mental disorders, reproductive disorders, antibiotic resistance, chest pain and interstitial pneumonitis which ultimately leads to death (Huang et al., 2019). It affects kidney, induces restlessness, and anxiety in exposed human population. Aquatic plants accumulate mercury inside them. Organic mercury is 200 times more lethal than inorganic mercury as it is formed by the combination of mercury and carbon resulting in the formation of compounds such as methylmercury which is neurotoxic in nature. In microbes, it causes protein denaturation, disruption in outer layer of cell, inhibit enzymatic activity as well as transcription (Salam et al., 2019). In plants, it alters the permeability of cell envelop and causes abnormal germination (Patra & Sharma, 2000). It decreases chlorophyll synthesis, photosynthesis, anti-oxidant defence mechanism and transpiration rate. It imbalances essential

Journal Pre-proof

ions like decreases potassium (K), magnesium (Mg), manganese (Mn) and increases iron (Fe) concentration (Salam et al., 2019).

2.5 Other heavy metals

Lead is used in various industries like coal, batteries, gasoline, pipelines and pigment. Zinc (Zn) is used in the production of pigments, metal alloys, pipelines and electroplating. Nickel is used in the production of batteries, vegetables and metal alloy (Chugh et al., 2020). These industries release their waste directly into the environment and pollute it. Naturally, copper is present in the ecosystem in the form of carbonates, sulfides and oxides. It is accumulated into the environment through the untreated effluent released by various industries like electroplating, mining, metal alloys and present in the domestic waste (Coelho et al., 2016). Cadmium (Cd), nickel (Ni), copper (Cu), and chromium (Cr) are mostly used in electroplating industries (Kapahi & Sachdeva, 2019). Textile industries release lead, copper, nickel, and chromium, mostly Cu is present in excess (Hassan & Carr, 2018). Nuclear power plants release Zn and Cu. Copper is used in the heat exchanger (Rajamohan et al., 2010). Lead is considered to be neurotoxic. Nickel causes skin allergy and affects the lungs and cardiovascular system. Zinc and copper lead to anemia (Coelho et al., 2016). Heavy metals like cadmium, lead, mercury and nickel inhibit cell division, protein and DNA denaturation. Arsenic, cadmium, nickel, lead, zinc, and copper leads to the inhibition of enzymatic activities and alter the cell membrane structure (Anyanwu et al., 2018).

3. Phycoremediation: an insight to the mechanism

Many heavy metals like molybdenum, copper, zinc, nickel, manganese, iron, cobalt and boron are considered to be micronutrients for microalgae. These trace elements helps in the growth of algal cells as they are essential for the cell metabolism (Sayed et al., 2019). On the other hand heavy metals like silver, gold, aluminum, mercury, titanium, cadmium, lead, and arsenic are not beneficial for microalgal growth and considered to be toxic heavy metals (Jais et al., 2017). Microalgae is considered to be outstanding bio-remediators because of their excellent tolerant capacity, easy to grow, high binding affinity, large surface area, eco-friendly, re-usable and dead biomass can be used for remediation i.e. no nutrient or other parameters are required and is labor extensive (Ramakrishnan et al., 2010) (Darda et al., 2019). They have multiple mechanisms for the self-protection against heavy metals like gene

Journal Pre-proof

regulation, chelation, and so forth. Desorption in algae is observed by reducing the pH with the help of hydrochloric or citric acid (Sayed et al., 2019).

Phycoremediation of heavy metals with the help of microalgae is a two-stage process. These techniques use bio-binding or bio-removal methods. In the first stage, heavy metals are adsorbed over the cell surface (biosorption) as shown in **fig.1**. It is a rapid process and may or may not involve cell metabolism. So, this stage is known as rapid extracellular passive process. Algae act as potential bio-sorbents because of the presence of different functional groups on the cell surface. Owing to their microscopic nature, microalgal cells are able to biosorb heavy metals very efficiently as they have a high surface area and high binding affinity. Algal cells possess the capability to bind up to 10% of its total biomass with metals. Pollutants bind to the cell surface through electrostatic force of attraction and complexation (Zohoorian et al., 2020). This process does not require any energy therefore, it is referred to as a passive process. Biosorption efficiency is governed by various factors such as bioavailability of metals, availability of metal binding groups on the algal cell surface, metal uptake and storage efficiency of algal cells. Biosorption can be carried out by both living and non-living biomass. Biosorption of algae involve two mechanisms. It can either exchange metal ions with calcium, magnesium, sodium or potassium ions which are present on the surface of algal cell or it can form complex with the functional groups which are present on the cell surface (Shamim, 2018). Various bio-sorbents are studied for the effective removal of contaminants. Biosorption is more economical as it can abate pollutants with the help of non-living biomass. Contaminants can easily bind to the surface of algae due to the presence of polysaccharides, lipids, and proteins. This result in the flocs formation and reduces the pollutants in the wastewater. The cell wall of microalgae consists of functional groups like sulfate, carboxyl, amino, and hydroxyl which are the binding site for the pollutants (Ahmed et al., 2021).

Fig.1 Two stage mechanism: remediation of heavy metals

In the second stage, heavy metals are transported across the cell membrane to cytoplasm or other organelles. This helps in the accumulation of the heavy metals inside the algal cells and this process is known as bioaccumulation (**fig 1**.). Bioaccumulation is an active process and thus, requires energy for the accumulation of pollutants inside the cell (Chugh et al., 2020). This process involves metabolic process for remediation thus, can be carried out by living biomass only. This process is also known as slow intracellular active accumulation because it is a slow process and

Journal Pre-proof

requires energy for the accumulation of heavy metals inside the microalgal cells. The algal biomass thus obtained with the help of these processes, can be further used for the recovery of some value-added products (Rezvani et al., 2018). After bioaccumulation, these pollutants are detoxified, removed or form a non-toxic complex with the help of phenomenon detoxification, compartmentalization or complexation.

3.1 Self-defense mechanisms of algae against heavy metal toxicity

Algae have special features of self-defense against the toxicity of heavy metals. They perform gene regulation, complexation, ion-exchange, chelation, produces reducing agents or anti-oxidants and cause heavy metal immobilization. These processes helps microalgae to fight against the toxicity of heavy metals which inhibit cellular metabolism, enzyme inhibition, abnormal physiological and morphological development, protein denaturation, nucleic acid, and cell membrane defects (Priyadarshini et al., 2019). These self-protection mechanisms help algae for their survival in the higher concentration of heavy metals present in the wastewater and helps in the abatement of toxic heavy metals. The self-defense mechanisms shown by algae are basically categorized into four main types which are biosorption, bioaccumulation and compartmentalization, biotransformation or metal detoxification and metal exclusion shown in **fig. 2.** In the following section we will discuss how these self-protection mechanisms helps in the removal of heavy metals from the wastewater.

Fig. 2 Self-defense mechanisms of algae against heavy metal toxicity

3.1.1 Extracellular uptake (biosorption)

It is an extracellular self-protection mechanism observed in microalgae that prevents the intake of heavy metals inside the algal cell. Adsorption of heavy metals occurs on the cell surface with the help of functional groups of polysaccharides, lipids, proteins, monomeric alcohols present on the algal surface. These compounds are rich in functional groups like hydroxyl, carboxyl, phosphate (Beni & Esmaeili, 2020). All these groups are anionic in nature due to which they attract positively charged metal ions. The binding of these functional groups with positively charged metal ions are due to the van-der waal, ion-exchange, covalent, or combination of these bonding (Tu et al., 2018). Every strain has different cellular composition; due to this very reason metal sorption varies with strain to strain. Wastewater containing heavy metals like copper, lead, zinc, and cadmium of concentration 40 ppm was

Journal Pre-proof

abated by freshwater and marine algae with the help of functional groups like hydroxyl, carboxyl, and amino groups present on the cell wall (Utomo et al., 2016).

Brown algae are observed to remediate maximum amount of heavy metals as compared to other algae. Cell wall of brown algae consists of functional groups like cellulose, alginic acid and other polymers. Brinza et al observed that brown algae *Fucus vesiculosus* is able to remove zinc with the help of carboxyl and hydroxyl groups present on the algal surface (Brinza et al., 2020). On the other hand, *Fucus vesiculosus* was able to remediate 143.2 ± 7.5 cadmium, 70.1 ± 1.9 nickel, 516.3 ± 12.5 lead mg g−1 respectively (Moreira et al., 2019). Similarly, Khajavian et al. also observed that brown algae *Cystoseria indicant* was able to abate 18.17 mg/g Nickel and 55.34 mg/g cadmium (Khajavian et al., 2019). Blue-green algae *Anbaena spharica* was demonstrated to remediate 111.1 mg/g Cadmium and 121.95 mg/g lead with the help of amino, hydroxyl and carboxyl functional groups present on the cell surface (Abdel -Aty et al., 2013). It is also observed that lower pH protonates the functional group which results the desorption of heavy metals and helps in the separation of various economical heavy metals (Sayed et al., 2019). On the other hand, alkaline pH favors the metal uptake. This process is highly efficient, as both living and non-living biomass can perform it, cost-effective and don't produce any toxic metabolites. Further, it doesn't affect the accumulation of co-product in algae as well as any cellular metabolism.

3.1.2 Intracellular uptake (bioaccumulation and compartmentalization)

After first self-defense mechanism against heavy metals which prevents the entry of metal ions inside the cell, heavy metals manage to enter the cells if present in higher concentration in the surrounding. To fight back the metal ions toxicity inside the microalgal cells, they show another mechanism of self-protection, termed as bioaccumulation or compartmentalization, which helps in the remediation of heavy metals. In bioaccumulation, metal ions are accumulated inside the microalgal cells. When these ions are accumulated inside the algal cell organelles (vacuoles or thylakoids) then this process is known as compartmentalization (Zohoorian et al., 2020). These processes are slow as whole process is time consuming from the uptake of metal ions to the transport of these ions inside the cell or any cell organelles. Transport of metal ions takes place with the help of ion-selective transport proteins present over the cell membrane (Arunakumara & Xuecheng, 2008). These processes require energy for the accumulation. This is a metabolism dependent process and can be performed by the living cell only. Thus, this is a slow intracellular active bioaccumulation self-protection strategy performed by microalgal cells. Due to their small size i.e. large surface to

Journal Pre-proof

volume ratio, these metal ions are easily uptake by the microalgal cells. It is observed that sulphur (S), iron (Fe) and nitrogen (N) helps in the bioaccumulation of the metal ions (Shanab et al., 2012). Upadhyay *et al* reported that *Chlorococcum* sp. accumulated 239.09 ug/g arsenic from the wastewater. This species acts as a potential bioindicator and bioremediator. Green algae *Tetraselmis suecica* was able to accumulate cadmium inside the vacuole and *Skeletonema costatum* can bioaccumulate cadmium and copper inside the vacuole (Perales-Vela et al., 2006). Massocato *et al* observed that manganese and nickel are taken up by the electron-dense bodies inside the *Ulothrix* sp. LAFIC 010 (Massocato et al., 2018).

3.1.3 Metal detoxification

Metal Detoxification or biotransformation is the third type of self-protection mechanism shown by microalgal cells where heavy metals are converted into non-toxic compounds with the help of chelators. Chelating agents are the organic compounds which link with the metal ions and form complexes. This process is also known as complexation. It is observed that algae releases chelating agents in defense of metal toxicity. These chelating agents binds with the metal ions and convert them into a non-toxic form (Genchi et al., 2020). These complexes of heavy metals with other compounds get easily accumulated inside the cell or other cell organelles (compartmentalization) or these non-toxic complexes are removed from the microalgal cell with the help of efflux transporters. Algae secretes many chelating agents like organic acids, peptides, enzymes, thiol containing molecules, and so forth (Leong & Chang, 2020) (Navakoudis & Ververidis, 2018). Peptides like phytochelatin or metallothioneins helps in the detoxification mechanism of microalgae. Perales-Vela et al observed that metallothioneins released by *Phaeodactylum tricornutum* was able to detoxify copper ions. *Phaeodactylum tricornutum* release metallothioneins and forms cadmium-metallothioneins complex with sulphide ions (Perales-Vela et al., 2006). Nowicka *et al* reported that reactive oxygen species, detoxifying enzymes and hydrophilic antioxidants are released by *Chlamydomonas reinhardtii* as a self-defense mechanism against cadmium and chromium toxicity (Nowicka et al., 2020).

3.1.4 Metal exclusion

In this, heavy metals are removed by the microalgal cells with efflux transport. This mechanism is another type of self-defense followed by the microalgal cell. This strategy maintains heavy metal concentration inside the cell. This active transport system in microalgae, helps to get rid of heavy metals when they are lethal for cell or non-toxic form

Journal Pre-proof

of metal ions (complexation). This exclusion transport occurs with the help of ion-selective transporters. It is also observed that algae alter the cell membrane to reduce the influx of heavy metal (Gaur & Rai, 2001) (Arunakumara & Xuecheng, 2008). This self-protection mechanism is common in other microbial cells like bacteria or fungi (Priyadarshini et al., 2019) but more studies are required in the algal efflux mechanism for heavy metals for better understanding.

4 Phycoremediation of heavy metals: Cadmium, Arsenic, Chromium, and Mercury

Various conventional methods (ion exchange, electrochemical treatment, osmosis, evaporation, and precipitation) were used to remove the heavy metals from the wastewater but they are not effective as well as are expensive as compared to the modern techniques (Fomina & Gadd, 2014). Bioremediation is a clean technology that results in effective cleaning of the environment along with the production of biomass (Coelho et al., 2016). This technique can remediate the pollutants when they are present in a low amount when other physicochemical methods failed, and biomass produced can be used in the production of some value-added product.

In bioremediation, biological agents like bacteria, algae, fungi are used for the abatement of heavy metals. In the statistics, it is observed that algae is able to remediate approximately 15.3 to 84.6% of heavy metals as compared to fungi and bacteria (Leong & Chang, 2020). This has various advantages over other bioremediation techniques as algae has high uptake capacity for heavy metals, can be regenerated, algal biomass can be used all around the year, does not produce any toxic metabolites and it is a cost-effective technique.

Algae abate heavy metals through the two-stage process. It adsorbs the metals over the cell surface in the first stage. These ions are adsorbed because of the presence of functional groups (thioether, carboxylic, hydroxyl, amide, sulfhydryl, and imidazole) on the cell wall (biosorption). The positively charged ions get attracted towards these negative ions on the cell wall and abate the heavy metals. In the second stage, over time these metal ions are transported into the cell slowly (bioaccumulation) (El-Sheekh et al., 2015) and further detoxification, efflux or compartmentalization takes place as discussed in earlier sections. In the following section remediation of four most toxic heavy metals are briefly discussed:

4.1 Cadmium

Journal Pre-proof

It is seen that different strains of algae have different effect towards cadmium toxicity. It decreases growth as well as photosynthetic efficiency and enhances the secretion of chelators which helps in the detoxification of cadmium. Algae secrete phytochelatins or enzymes as a self-protection mechanism against toxicity of cadmium. Deng et al observed that low concentration of cadmium has minimal effect on the growth of *Microcystis aeruginosa*. However, higher concentration of cadmium secrete more malondialdehyde content and enzymes like esterase, mutase which inhibits growth as well as photosynthetic activity of the algal cell (Deng et al., 2020). Out of two green algae (*Scenedesmus accuminatus var biseratus* and *Scenedesmus protuberans*) and diatom *Cyclotella* sp., diatom show maximum tolerance as well as remediation of cadmium (99%) (M.-T. Vo et al., 2020). It might be possible only if maximum adsorbtion of cadmium is done by diatom as compared to bioaccumulation as if cadmium is accumulated inside the cell, it will decrease the photosynthetic activity as well as growth of the cell as happened in the two green algae. Chandrashekharaiah *et al,* observed cadmium abatement capacity of two freshwater algae *Scenedesmus acutus* and *Chlorella pyrenoidosa*. *Scenedesmus acutus* show maximum removal of 57.14% of cadmium where 98.5% is adsorbed over the surface of the cell and only 1.5% is accumulated inside the algal cell (Chandrashekharaiah et al., 2021). Similarly, marine red algae *Sarcodia suidae* was able to bioadsorb more as compared to bioaccumulation. Han et al observed the ratio of biosorption/ bioaccumulation rate in the dark and light phase for 1 and 5mg/L concentration of cadmium. They observed maximum abatement in the light phase for both the concentrations. In the light phase, biosorption/bioaccumulation ratio is 1.75 and 1.74 for 1mg/L and 5mg/L respectively. They also observed that as the bioaccumulation rate increases with time, algae tolerance decreases as chl-a concentration, photosynthetic and respiratory efficiency is reduced (Han et al., 2020). For the enhancement of the removal efficiency various innovative techniques are used for the modification of the biomass which leads to increase in the removal efficiency which is discussed in the recent advancement sections. It can be concluded that cadmium abatement capacity depends upon the strain of the bioremediators, mechanism of remediation, manipulation technique and on the removal parameters. More recent studies along with their removal parameters are summarized in **Table 1**.

Table 1 Phycoremediation of heavy metals along with the removal efficiency, time, mechanism, and growth parameters

Journal Pre-proof

4.2 Arsenic

Arsenic is one of the most toxic heavy metals as per US EPA guidelines. Physical and chemical properties of arsenic depend on their oxidation state i.e., arsenic, arsine, arsenite and arsenate. The most toxic form of arsenic is inorganic form. In phycoremediation, it is generally seen that microalgae try to detoxify arsenic by converting inorganic form into organic form. It can be converted with the help of chelators or functional groups like phosphate, nitrate (Hussain et al., 2021). Arsenic is abated by Chlorophytes (*Chlorella salina, Ostreococcus tauri, Dunaliella salina, Polyphysapeniculus, Dunaliella tertiolecta*); Heterokontophytes (*Heterosigma akashiwo, Phaeodactylum tricornutum, Skeletonema costatum, Thalassiosira pseudonana*) reported by Duncan, Maher and Foster (Duncan et al., 2015). Saavedra *et al* reported the removal efficiency of arsenic by green algae *Chlamydomonas reinhardtii, Chlorella vulgaris, Scenedesmus almeriensis* and an indigenous *Chlorophyceae* sp. Out of these four, *Scenedesmus almeriensis* was able to remediate 40.7% of arsenic within three hours. It was also observed that in multimetallic solution, removal efficiency of arsenic is increased by 1.7 factor (Saavedra et al., 2018). To enhance the removal efficiency of arsenic by *Chlamydomonas reinhardtii*, Ramírez-Rodríguez and co-workers modify the acr3 gene of *Chlamydomonas reinhardtii* and observed that abatement capacity of the acr3 modified gene is enhanced by the factor of 1.5 to 3 for different arsenic concentration as compared to wild strain (Ramírez-Rodríguez et al., 2019). On the other hand, other researchers recently tried to co-relate the effect of different parameters with the removal efficiency. Libatique *et al* exposed red algae, *Sarcodia suiae* to different temperature ranges (15,20 and 25°C) and observed total arsenic and arsenite are abated at lower temperature (15°C). Light have no such significant effect on the removal efficiency but concentration strongly effects the accumulation rate (Libatique et al., 2020). It can be concluded that removal of different oxidation state of arsenic is dependent on temperature variation as well as initial arsenic concentration. Henriques et al stated that salinity have no effect on the removal efficiency of arsenic in the case of living seaweed. They worked on *Ulva lactuca* and reported the 48% of abatement capacity within 72 hours (Henriques et al., 2019). **Table 1** summarizes the recent studies of remediation of arsenic by different algal strains along with other removal parameters.

4.3 Chromium

Chromium is mainly present in the two forms in environment i.e., trivalent chromium and hexavalent chromium. Hexavalent chromium can easily diffuse the biological membrane and then with the help of reduction mechanism

hexavalent form is reduced to trivalent form (Aharchaou et al., 2017). Chromium removal is done by various algae like *Cladophora, Spirogyra, Chlorella vulgaris, Nostoc linckia, Euglena, Ceramium virgatum* and *Spirulina* sp. It was observed that *Chlorella* can remediate upto 97% of chromium within 28 days. *Spirulina platensis, Pelvetia canaliculata* and *Sargassum filipendula* are able to remediate chromium from the contaminated water (Fernández et al., 2018). Different kinds of chromium removal techniques from algae are observed by the researchers. Biosorption, bioaccumulation and detoxification are the major mechanisms followed by algae for the abatement of chromium. Guarín-Romero et al reported that *Durvillaea antartica* was able to adsorb 102.72mg/g chromium over the functional groups (hydroxyl, amino, methyl and carboxylic) present on the extracellular surface of the brown algae at pH 5 within seven hours of time period (Guarín-Romero et al., 2019). Chromate and dichromate (hexavelant chromium) were adsorbed by two marine algal strains *Phaeodactylum tricornutum* strain CCY0033 and *Navicula pelliculosa* strain CCMP543. Both the strains were able to remediate 30-35% of hexavalent chromium (Hedayatkhah et al., 2018). On the other hand, Aharchaou and other researchers studied the efficiency of two most common mechanisms (biosorption and bioaccumulation). They observed that green algae *Chlamydomonas reinhardtii* was able to adsorb more trivalent chromium in the initial hours as compared to accumulation rate. Further they saw that rate of biosorption as well as bioaccumulation is same for the hexavalent chromium. Initially, they found that within 1 hour major portion of the chromium is adsorbed over the surface of algae but with time after 72 hours, accumulation rate increases within the cell with 37±7% in the organelles and 39±18% over the peptides or proteins (Aharchaou et al., 2017). Another way to remediation of chromium from the environment is through the algal enzymatic route. Lee et al reported that *Chlorella vulgaris* was able to completely remediate 10mg/L hexavalent chromium with the help of chromate reductase within 7 days (L. Lee et al., 2017). On the other hand, non-enzymatic pathway studied by Yen et al observed that dead biomass is also able to reduce hexavalent chromium into trivalent, might be because of the release of glutathione from the broken dead cells (Yen et al., 2017). For the enhancement of abatement capacity, environmentalists are using more advance techniques like pre-treatment with sodium hydroxide or sodium dodecyl sulphate or methylation, immobilization or development of biochar. All these advance techniques help in the complete removal of chromium as discussed in the later section of this article. More studies of heavy metal removal are added in the T**able 1.**

4.4 Mercury

It is one of the most toxic heavy metals, even its converted by-products are considered to be hazardous for the ecosystem. Algae are considered to be the good adsorber of mercury because of the presence of surface-active functional groups. Solisio and co-workers observed interaction between mercury and extracellular functional groups present over the surface of *Chlorella vulgaris* and stated that this is the electrostatic interaction due to the presence of oxygen lone pairs. *Chlorella vulgaris* is able to adsorbed 17.49 mg/g mercury with the help of carboxylic groups present over the surface of algae at pH 5, 20°C within 2 hours (Solisio et al., 2019). Huang *et al.* used *Scenedesmus obtusus* XJ-15 for the adsorption of mercury by the hydroxyl, carboxyl and phosphate containing functional groups. They further enhanced the removal efficiency by treating the algal biomass with phosphate and achieved 95 mg/g of sorption capacity at pH 5, 25°C within 180 minutes (Huang et al., 2019). Other than extracellular removal mechanism, enzymatic reduction of Hg^{2+} into zero-valent mercury with the help of mercuric reductase is also used for the abatement. Now this Hg^0 is volatilized with the help of biological (diatoms or green algae) or non-biological volatilization (light) methods. It is observed that biological volatilization is rapid (30 mins to few hours) than non-biological methods. This volatilization can be further enhanced by using chemical agents or transgenics. Bioaccumulation of mercury, induces phytochelators (thiol, methyl groups) and help in the removal of mercury (Leong & Chang, 2020). More advanced processes of mercury uptake are discussed later in the recent advancement section. More studies of mercury removal are mentioned in the **Table 1**.

4.5 Other heavy metals

All other toxic heavy metals like lead, zinc, uranium, nickel, silver, copper, and manganese etc, are remediated with the help of above-mentioned mechanisms only. For example, *Cladophora fascicularis, Fucus vesiculosus, Ascophyllum nodosum, Hydrodictylon, Oedogonium, and Rhizoclonium* sp., *Spirogyra* sp., *Spirulina* sp., *Chlorella pyrenoidosa*, etc., are the few examples of microalgae that helps in the remediation of heavy metals (Coelho et al., 2016). Ar, Zn, Cu, B and Mn are reported to be remediated by *Chlamydomonas reinhardtii*, *Chlorella vulgaris, Scenedesmus almeriensis* and *Chlorophyceae* sp. It is observed that manganese is abated by *Chlorella vulgaris*. It can remove 99.4% of Mn within three hours at 7 pH. The removal efficiency of copper is 88% achieved by *Chlorophyceae* sp. at pH 7 within 10 minutes. *Scenedesmus almeriensisis* can remove 38.6% of boron within 10 minutes at pH 5.5. 91.9% of zinc is remediated by *Chlorophyceae* sp. within three hours at pH 5.5 (Saavedra et al.,

Journal Pre-proof

2018). *Chlorella pyrenoidosa* is able to absorb copper and lead from the environment (Shi et al., 2017). More studies of different heavy metal removal are added in the T**able 1**.

5 Phycoremediation: a sustainable remediation technique

As discussed in the above section, phycoremediation has great potential for the treatment of wastewater. This process is more eco-friendly as compared to the conventional treatment, but it is a cost-prohibitive technique as algae cultivation needs to be done in optimal conditions. The maintenance of the optimal growth conditions for the maximum rate of biomass and removal efficiency makes this process more expensive. To make this process not so cost-prohibitive, integrated systems are designed (Kaloudas et al., 2021). In these integrated systems, two or more processes are combined to make this process more sustainable (Nur & Buma, 2019). In these systems, one can have a hybrid technique where wastewater treatment plant is combined with the production of some innovative by-products from the biomass as shown in **fig. 3**. These innovative products can be enzymes, bioenergy, biopolymers, pigments, and sugars, etc., which can be further used by the society. Thus, it is a promising technique for industries as well as for the environment (Emparan et al., 2019). At the same time, it sequesters the pollutants and gives rise to pollution-free water bodies with some commercial products (Quijano et al., 2017) (Koreiviene et al., 2019).

Fig. 3 Economical set-up for the remediation of industrial wastewater

These innovative methods are used where bioremediation is carried out. Microalgae have different benefits like it has high carbohydrate, protein, and lipid content which can be used for the production of various new commercial as well as an economical products like fodder, remediation, organic fertilizer, green energy, bioenergy, nanoparticles, organic fertilizer, cosmetic and pharmaceuticals (Kumar & Bharadvaja, 2020a) (Ravindran et al., 2016) (Enamala et al., 2018). Different value-added products produced by the algae are mentioned in **Table 2** along with the heavy metals abated by respective algae.

Table 2 Different value-added products along with the heavy metals abated by respective algae

Journal Pre-proof

5.1 Phycoremediation of heavy metals along with the production of Bioenergy

Microalgae take its nutrients from waste and give rise to biomass which is very economical as well as can be used for the treatment (D. V. Singh et al., 2021). Various algae like cyanobacteria are known to have higher levels of lipid in it, which can be used to produce biodiesel. It is believed that algae are the only eco-friendly source for the production of biodiesel. Algae-based biofuel is referred to as green energy and it leads to the production of the economical product (H. N. P. Vo et al., 2019) (Vassilev & Vassileva, 2016). *Dunaliella tertiolecta, Spirulina maxima, Chlorella vulgaris, Neochlorisoleo abundans* and *Scenedesmus obliquus* are used for the production of bioenergy because of the presence of higher levels of oil in it. Various strains of algae are used in this hybrid technique are *Chlamydomonas mexicana, Nitzschia cf. pusilla, Chlorella vulgaris, Micractinium ereisseri, Ourococcus multisporus* and *Scenedesmus obliques* (Baghour, 2019). Algae are responsible for 40% of the global carbon fixation. Seaweeds are one of the best algae for the production of bioenergy because of the high sequestration rate of CO_2, show maximum growth as well as production rate (Tabassum et al., 2017). In this subsection we will discuss remediation of different heavy metals and how these heavy metals or more advance technologies are enhancing the accumulation of lipid as well as remediation of heavy metals. Positive influence of different heavy metals is presented in **fig. 4**. It is observed that cadmium, copper and lead leads to increase the phytohormones in algae. Similarly, arsenic, cadmium,, copper, iron and nickels helps in the production of lipids and pigments. Arsenic, cadmoim, copper and lead leads to the increase in amount of phytochelatin and silver, cadmium, cobalt and copper show positive influence on exopolymers.

Fig. 4 Heavy metals enhancing the accumulation of by-products

As heavy metals are toxic to the algae, they tend to induce oxidative stress in them. It was observed by Upadhyay *et. al.,* that this oxidative stress increases the lipid content in the microalgae which is beneficial for the production of biofuel (A. K. Upadhyay et al., 2016). Yang *et. al.,* stated that *Chlorella minutissima* UTEX2341 was able to remediate 62.05% of zinc, 83.68% of manganese, 74.34% of cadmium and 83.60% of copper from the wastewater and it was further observed that cadmium and copper enhanced the production of lipid content in the range 21.07%-93.90% (Yang et al., 2015) whereas Zhao et al showed 49.37% increase of lipid content with the help of gamma-aminobutyric acid in *Monoraphidium sp.* QLY-1 in the presence of cadmium (Zhao et al., 2019).

Journal Pre-proof

Heterochlorella sp. MAS3 and *Desmodesmus sp*. MAS1 can abate cadmium more than 58% and bioaccumulate lipids in hydrocarbons and triglycerides at pH 3.5 (Abinandan et al., 2019).

Hedayatkhah *et. al.*, reported that hexavalent chromium enhances the lipid production by 43 to 300% in marine algae (*Phaeodactylum tricornutum* and *Navicula pelliculosa*) at temperature 23°C and 27°C respectively. They further showed the enhancement of lipid accumulation depends on the nitrogen source, light intensity and observed that 3mM nitrogen source and 50 umole $m^{-2}s^{-1}$ gives the best result (Hedayatkhah et al., 2018). Another more advanced system which enhances the heavy metal abatement along with lipid production is biomimetic mineralization. Here polyelectrolyte layer is formed over the green algae *Chlorella vulgaris* where calcium ions are completely trapped. This enhances the mercury removal from 62.85% to 94.74% and also increase the desorption capacity from 48.42% to 88.72% as compared to raw algal cells. Further, it is also observed that same lipid production of 9.17% is achieved which was reduced to 5.12% earlier in the raw cells when exposed to mercury (Peng et al., 2017).

6 Recent advancements in phycoremediation approach towards heavy metals removal from wastewater

Microalgal based bioremediation is carried out for decades now. However, with time, to make the process more efficient new technologies are explored and incorporated. Recent studies have focussed on remediation of heavy metals with the help of microbial consortia instead of using a single bioremediatory agent. In other experiments, it is observed that immobilization of biomass also enhances the sorption capacity of the heavy metals. In this section, methods like hybrid systems, innovative reactors, polymer or organic compounds, and modification of biomass or transgenic based enhancement of heavy metal remediation is discussed.

It is shown that consortia of algae and other remediators are able to show more removal efficiency that to in a small period of time as compared to the single bioremediatory agent (Joutey et al., 2014) (Qian et al., 2016). It was observed that algae and fungi or bacteria interactions lead to the flocculation of the algae under unfavourable environment (Siddeeg et al., 2020)(Quijano et al., 2017). Thus, this technique reduces the harvesting activities of algae (Lavrinovičs & Juhna, 2018). Green algae consortia with bacteria like cyanobacterial species, *Dunaliella, Clostridium dunii, Chlorella vulgaris, Clostridium needles* and *Chlamydomonas salina* was used for the abatement of arsenic. Arsenic was adsorbed by the cell wall of algae through functional groups like , hydroxyl, carbonyl and thiol which are present on the surface of the cell wall (Hussain et al., 2021). Gao et al. reported that 4.8×10^7 cells

Journal Pre-proof

mL^{-1} algal-fungal of *Chlorella vulgaris* and *Aspergillus oryzae* biomass at 140 rpm, pH 7 and 5gm/L glucose concentration show maximum rate of remediation of arsenic (Z. Gao et al., 2020). Bodin et al reported that biopellets of green algae C*hlorella vulgaris* and fungi *A. niger* is able to enhance cadmium remediation from 40 to 56% as compared to *Chlorella vulgaris* alone (Bodin & Hultberg, 2017).

The stability of the process and the removal efficiency are enhanced by the immobilization of the microbial cells. It can also be enhanced by combining the cells or enzymes with nanotechnology (Fernández et al., 2018)(Borji et al., 2020). Maintenance of the algae and for effective removal, immobilization of the algae on alginate beads is carried out in the bioreactor where wastewater is treated (F. Gao et al., 2019) (Luo et al., 2017). Immobilized *Gelidium amansii* was able to remediate 100% lead of concentration 200mg/l within three hours (El-Naggar et al., 2018). Huang *et al* reported that *Scenedesmus obtusus* was immobilized for the removal of mercury. First, microalgae are surface modified by growing under different phosphate concentrations and then they are immobilized. This prevents the leaching of phosphate as well as helps in the immobilization of mercury. It enhances the sorption capacity of mercury. This strain will remediate the phosphate and mercury from the contaminated water (Huang et al., 2019). *Cladophora* spp. was immobilized on alginate beads which were used for the adsorption of mercury. It is observed that hydroxyl, carboxyl and amine groups are used in the interaction of algal cell with mercury and maximum removal is achieved at pH 6, 16°C within an hour (Mokone et al., 2018).

Another innovative method that is recently being used for bioremediation is using microbial biochar in the form of carbon dots. They not only show bioremediation but also detect the presence of heavy metals such as lead, nickel, copper and cadmium (Plácido et al., 2019). Nowadays, researchers are extensively studying the carbon-based material for the adsorption of different heavy metals. Biochar is obtained after pyrolysis in the absence of oxygen this material is re-usable, cost-effective as well as eco-friendly removal technique (Lebrun et al., 2021). Senthilkumar *et al* recently abated arsenic with the help of biochar obtained from green algae, *Ulva reticulata*. They were able to adsorbed 8.12 mg/g arsenic(V) and regenerated this material with the help of sodium hydroxide as it show desorption efficiency as 99.5% (Senthilkumar et al., 2020).

Recently Wei et al. recorded that synthetic organic pollutants increase the removal efficiency of heavy metal by the bioremediator. They combine o-nitrophenol with chromium. It was observed that this combination decreases the reactive oxygen species and increases cell damage. Thus, now these cells can be reused approximately for 4 times.

They observed that *Chlamydomonas reinhardtii* remediation rates of chromium was increased from 37.4 % to 54.9% and 35.8% to 45.9% for o-nitrophenol (Wei et al., 2020). Wu, Ma and Zeng observed that extracellular polymeric substances increase the removal capacity for heavy metals like chromium by using iron/carbon-based nano-composite. It can absorb 16.56% and remediate 83.44% of chromium (J. Wu et al., 2018). Radwan et al. reported that *Chlorella vulgaris* biomass was heat treated and modified with citric acid and sodium hydroxide is able to abate 92% copper ions (Radwan et al., 2020). Zhang *et al* compared the uptake of arsenic in extracellular polymeric substances (EPS) covered *Chlorella pyrenoidosa* algal cell as compared to EPS free algal cell. It was observed that EPS-algal cell was able to adsorb more arsenic and reduces intracellular accumulation. For the first time, in 2019 Zhang and co-workers are able to identify the functional groups (C-O-C, C-O-H and NH_2) which link the arsenic with the tyrosine and tryptophan like proteins on EPS (Zhang et al., 2019). Dead biomass of *Phaeodactylum tricornutum* and *Navicula pelliculosa* were able to remediate 24-32% hexavalent chromium but this efficiency was further enhanced in the presence of extracellular polymeric substances. EPS covered algal cells show 27-37% abatement efficiency of chromium (Hedayatkhah et al., 2018). Moreover, Naveed *et al.,* reviewed the role of microalgal EPS for the sequestration of metalloids (Naveed et al., 2019).

Recently, Costa *et. al.,* designed a photocatalytic hybrid system that consists of titanium dioxide immobilized in the alumina hollow fibres. *Chlorella vulgaris* was coated on the outer layer of the fibre. This hybrid system can remove chromium heavy metal from the environment and is reusable (Costa et al., 2019). Hybrid systems are used for making the bioremediation process continuous. Urea formaldehyde along with *Spirogyra* is together used for the making of granule-packed column reactor which can be used for the abatement of copper and lead up to 6-9 cycles. Granules diameter varies from the range of 0.2 to 0.5 mm. With the help of these granules, desorption of heavy metals is done easily and biomass can be further used for economical purposes (A. Singh et al., 2012). Recently, for higher removal efficiency, microbial fuel cells were used. This system consists of a single chamber and is made up of biocathodes. This helps abatement of heavy metals, color and chemical oxygen demand and helps in the generation of electricity. Logroño et. al., observed that a higher removal rate of zinc and chemical oxygen demand was seen as compared to chromium (Logroño et al., 2017).

7 **Challenges and Opportunities**

Algae are considered to be an outstanding bioremediator because of their excellent metal tolerant capacity, easy to grow, high metal binding affinity, large surface area, and eco-friendly nature. Even the dead algal cells can be used for remediation i.e. no nutrient or other parameters are limitations to its applicability for metals removal. But when we talk about phycoremediation of heavy metals there are some challenges which are faced by the scientists. During phycoremediation of heavy metals, despite metal tolerance capacity of algae, productivity of the biomass decreases and so decreases the productivity of lipids, nanoparticles, pigments and other valuable products. To overcome this more advance systems like biomimetic system or pre-treatment technologies should be used (Leong & Chang, 2020). Strains should be selected not only for maximum remediation of heavy metal but also able to accumulate maximum by-products like lipids, proteins, pigments, etc. To make the phycoremediation process more efficient, it is very important to carry out studies on the gene level and modify the genetic make-up of the microalgae to enhance the efficacy of the process (Darda et al., 2019). With the complete genomic studies, it will be possible to make the biomarkers for the contaminated site. CRISPR-Cas 9 tool can be used to make this process more efficient (Scranton et al., 2017b). Toxicological studies of the intermediates produced after transforming the pollutant should be done along with the characterization of the strains.

Alone bioremediation is not an economical process. Algae are harvested on wastewater open system, this results in reducing the biomass as algae has to share its nutrients with other biological agents like bacteria, fungi or pathogens which are naturally present in the wastewater. This can be overcome by using algal consortia with other biological agents which not only decreases the remediation time but also enhances the remediation efficiency (C. S. Lee et al., 2016). Another problem, which is faced nowadays, is that these bioremediators are grown in the waste stream just to make this process cost effective but it is observed that this wastewater contains some nutrient in excess or other in minute amount which effects the growth of the algae. So, it is very important to characterize the wastewater and then balance of the nutrients should be done. Another way is that one can isolate and screen the algal bioremediator from that waste stream. Pre-treatment techniques like membrane filtration, acidification, chlorination, UV irradiation can be performed to overcome these challenges (Z. Z. Wu et al., 2018) (Leong & Chang, 2020).

Till now complete remediation of heavy metals is not seen much often. This can be overcome by using genetic engineering, metabolic engineering or protein engineering by which alterations in the genes, proteins or metabolic pathway can be made which leads to the overexpression of protein used for the remediation. By using transgenics,

Journal Pre-proof

one can also enhance the productivity of the by-product like pigments, nutraceuticals, vitamins, and proteins, etc. After the remediation, one of the main problems which is faced by the scientists is to extract heavy metals from the bioremediators because they have poor desorption efficiency (Pratush, 2018). This can be enhanced by using much more advance techniques where modification of biomass is done with the help of extracellular polymeric substances, alginate and other chemicals (Naveed et al., 2019). This not only enhances the biosorption capacity but also enhances the desorption efficiency. With the help of this, scientists prevent the release of undesired heavy metals into the environment. All these strategies will make the microalgae as a promising tool for remediation of heavy metals by making it more efficient, cost-effective and more environment friendly. More studies at the molecular level are needed in the field of tolerance of heavy metals by algae, resistance mechanism carried out by algae. Further studies are required to make it a technology which simultaneously detects as well as abate the heavy metals from the environment. Still, a lot needs to be understood in the multi-disciplinary field to make this process a powerful technique for remediation of heavy metals in the large scale.

8 Conclusion

Phycoremediation is the most sustainable technology as compared to conventional technologies used for the treatment of wastewater. Heavy metals are non-biodegradable in nature and are hazardous for humans as well as for the flora and fauna. These pollutants are released directly or indirectly by several industries like electroplating, dyeing, batteries and automobile into various waterbodies. Algae are the hyperaccumulator of the heavy metals and are the source of O_2 in the aquatic atmosphere. These species produce a large amount of biomass which can be further used for the remediation and to produce innovative products like biodiesel, biomethane, organic fertilizers, feedstock, nanoparticles and pharmaceutical products. It is seen that various heavy metals remediation varied with strains as different strains have different physical and metabolic properties. Heavy metals are majorly abated with the help of two mechanisms which are biosorption and bioaccumulation. After bioaccumulation, toxic heavy metals are converted into non-toxic by using detoxification mechanism and transported out of the cell by using efflux mechanism. Thus, processes like adsorption, accumulation, complexation, compartmentalization, chelation, reduction play an important role in remediation of heavy metals. Recently, more studies are carried out in the area where more economical and sustainable approach of phycoremediation is experimented. Current research focuses

more on the process sustainability and generation of commercially important by-products. In these systems, one can have a hybrid technique where a wastewater treatment plant is combined with the production of some innovative products from biomass. Further, recent advancements like immobilization, algal consortia, modification of biomass, innovative and hybrid systems are discussed in this review for the enhancement of the removal efficiency of the microalgae. Still, more studies are needed in the molecular level to fully explore the heavy metal resistance mechanism of algae, to enhance the rate of remediation as well as production of by-product.

Acknowledgments: The authors would like to thank Delhi Technological University, Delhi, India for providing facilities for the output in this paper form.

Funding

None.

Disclosure Statement

The authors declare that they have no competing interest.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statements

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Journal Pre-proof

Journal Pre-proof

References

Abdel -Aty, A. M., Ammar, N. S., Abdel Ghafar, H. H., & Ali, R. K. (2013). Biosorption of cadmium and lead from aqueous solution by fresh water alga Anabaena sphaerica biomass. *Journal of Advanced Research*, *4*(4), 367–374. https://doi.org/10.1016/j.jare.2012.07.004

Abinandan, S., Subashchandrabose, S. R., Venkateswarlu, K., Perera, I. A., & Megharaj, M. (2019). Acid-tolerant microalgae can withstand higher concentrations of invasive cadmium and produce sustainable biomass and biodiesel at pH 3.5. *Bioresource Technology*, *281*(January), 469–473. https://doi.org/10.1016/j.biortech.2019.03.001

Aharchaou, I., Rosabal, M., Liu, F., Battaglia, E., Vignati, D. A. L., & Fortin, C. (2017). Bioaccumulation and subcellular partitioning of Cr(III) and Cr(VI) in the freshwater green alga Chlamydomonas reinhardtii. *Aquatic Toxicology*, *182*, 49–57. https://doi.org/10.1016/j.aquatox.2016.11.004

Ahmed, D. A. E. A., Gheda, S. F., & Ismail, G. A. (2021). Efficacy of two seaweeds dry mass in bioremediation of heavy metal polluted soil and growth of radish (Raphanus sativus L.) plant. *Environmental Science and Pollution Research*, *28*(10), 12831–12846. https://doi.org/10.1007/s11356-020-11289-8

Al-Homaidan, A. A., Al-Qahtani, H. S., Al-Ghanayem, A. A., Ameen, F., & Ibraheem, I. B. M. (2018). Potential use of green algae as a biosorbent for hexavalent chromium removal from aqueous solutions. *Saudi Journal of Biological Sciences*, *25*(8), 1733–1738. https://doi.org/10.1016/j.sjbs.2018.07.011

Ali, H., Khan, E., & Ilahi, I. (2019). Environmental chemistry and ecotoxicology of hazardous heavy metals: Environmental persistence, toxicity, and bioaccumulation. *Journal of Chemistry*, *2019*. https://doi.org/10.1155/2019/6730305

Ali Redha, A. (2020). Removal of heavy metals from aqueous media by biosorption. *Arab Journal of Basic and Applied Sciences*, *27*(1), 183–193. https://doi.org/10.1080/25765299.2020.1756177

Anyanwu, B. O., Ezejiofor, A. N., Igweze, Z. N., & Orisakwe, O. E. (2018). Heavy metal mixture exposure and effects in developing nations: An update. *Toxics*, *6*(4), 1–32. https://doi.org/10.3390/toxics6040065

Arora, N., Gulati, K., Patel, A., Pruthi, P. A., Poluri, K. M., & Pruthi, V. (2017). A hybrid approach integrating arsenic detoxification with biodiesel production using oleaginous microalgae. *Algal Research*, *24*, 29–39. https://doi.org/10.1016/j.algal.2017.03.012

Arunakumara, K. K. I. U., & Xuecheng, Z. (2008). Heavy Metal Bioaccumulation and Toxicity with Special Reference to Microalgae. *J. Ocean Univ. Chin. (Oceanic and Coastal Sea Research)*, *7*(1), 60–64. https://doi.org/10.1007/s11802-008-0060-y

Baghour, M. (2019). Algal degradation of organic pollutants. *Handbook of Ecomaterials*, *1*, 565–586. https://doi.org/10.1007/978-3-319-68255-6_86

Beni, A. A., & Esmaeili, A. (2020). Biosorption, an efficient method for removing heavy metals from industrial effluents: A Review. *Environmental Technology and Innovation*, *17*, 1–34. https://doi.org/10.1016/j.eti.2019.100503

Billah, M. M., Mustafa Kamal, A. H., Idris, M. H., & Ismail, J. (2017). Mangrove Macroalgae as Biomonitors of Heavy Metal Contamination in a Tropical Estuary, Malaysia. *Water, Air, and Soil Pollution*, *228*(9), 1–14. https://doi.org/10.1007/s11270-017-3500-8

Bodin, H., & Hultberg, M. (2017). Effects of biopellets composed of microalgae and fungi on cadmium present at environmentally relevant levels in water Hristina. *Int. Journal Phytorem*, *19*(5), 1–19. https://doi.org/https://doi.org/10.1080/15226514.2016.1244170

Borah, D., Kennedy, B., Gopalakrishnan, S., Chithonirai, A., & Nooruddin, T. (2020). Bioremediation and Biomass Production with the Green Microalga Chlorococcum humicola and Textile Mill Effluent (TE). *Proceedings of the National Academy of Sciences India Section B - Biological Sciences*, *90*(2), 415–423. https://doi.org/10.1007/s40011-019-01112-x

Borji, H., Ayoub, G. M., Bilbeisi, R., Nassar, N., & Malaeb, L. (2020). How Effective Are Nanomaterials for the Removal of Heavy Metals from Water and Wastewater? In *Water, Air, and Soil Pollution* (Vol. 231, Issue 7, pp. 1–35). Springer. https://doi.org/10.1007/s11270-020-04681-0

Brinza, L., Geraki, K., Cojocaru, C., Holdt, S. L., & Neamtu, M. (2020). Baltic Fucus vesiculosus as potential bio-

Journal Pre-proof

sorbent for Zn removal: Mechanism insight. *Chemosphere*, *238*, 1–11. https://doi.org/10.1016/j.chemosphere.2019.124652

Buayam, N., Davey, M. P., Smith, A. G., & Pumas, C. (2019). Effects of copper and pH on the growth and physiology of Desmodesmus sp. AARLG074. *Metabolites*, *9*(5), 1–17. https://doi.org/10.3390/metabo9050084

Chandrashekharaiah, P. ., Sanyal, D., Dasgupta, S., & Banik, A. (2021). Cadmium biosorption and biomass production by two freshwater microalgae Scenedesmus acutus and Chlorella pyrenoidosa: An integrated approach. *Chemosphere*, *269*(xxxx), 1–12. https://doi.org/10.1016/j.chemosphere.2020.128755

Chaudhary, R., Nawaz, K., Khan, A. K., Hano, C., Abbasi, B. H., & Anjum, S. (2020). An overview of the algae-mediated biosynthesis of nanoparticles and their biomedical applications. *Biomolecules*, *10*(11), 1–36. https://doi.org/10.3390/biom10111498

Chugh, M., Kumar, L., & Bharadvaja, N. (2020). Bioremediation of Heavy Metals: A step towards environmental sustainanbility. In *India 2020: Environmental Challenges, policies and green Technology.* (Vol. 66, pp. 77–90).

Chugh, M., Kumar, L., Bhardwaj, D., & Bharadvaja, N. (2022). Bioaccumulation and detoxification of heavy metals: an insight into the mechanism. *Development in Wastewater Treatment Research and Processes*, 243–264. https://doi.org/10.1016/B978-0-323-85657-7.00013-4

Coelho, L. M., Rezende, H. C., Coelho, L. M., de Sousa, P. A. ., Melo, D. F. O., & Coelho, N. M. M. (2016). Bioremediation of polluted waters using microorganisms. In *IntechOpen* (pp. 1–13).

Costa, I. G. F., Terra, N. M., Cardoso, V. L., Batista, F. R. X., & Reis, M. H. M. (2019). Photoreduction of chromium(VI) in microstructured ceramic hollow fibers impregnated with titanium dioxide and coated with green algae Chlorella vulgaris. *Journal of Hazardous Materials*, *379*(June), 1–10. https://doi.org/10.1016/j.jhazmat.2019.120837

Darda, S., Papalas, T., & Zabaniotou, A. (2019). Biofuels journey in Europe: Currently the way to low carbon economy sustainability is still a challenge. *Journal of Cleaner Production*, *208*, 575–588.

Journal Pre-proof

https://doi.org/10.1016/j.jclepro.2018.10.147

Deng, J., Fu, D., Hu, W., Lu, X., Wu, Y., & Bryan, H. (2020). Physiological responses and accumulation ability of Microcystis aeruginosa to zinc and cadmium: Implications for bioremediation of heavy metal pollution. *Bioresource Technology*, *303*(November 2019), 1–11. https://doi.org/10.1016/j.biortech.2020.122963

Duncan, E. G., Maher, W. A., & Foster, S. D. (2015). Contribution of arsenic species in unicellular algae to the cycling of arsenic in marine ecosystems. *Environmental Science and Technology*, *49*(1), 33–50. https://doi.org/10.1021/es504074z

El-Naggar, N. E. A., Hamouda, R. A., Mousa, I. E., Abdel-Hamid, M. S., & Rabei, N. H. (2018). Biosorption optimization, characterization, immobilization and application of Gelidium amansii biomass for complete Pb2+ removal from aqueous solutions. *Scientific Reports*, *8*(1), 1–19. https://doi.org/10.1038/s41598-018-31660-7

El-Sheekh, M. M., Farghl, A. A., Galal, H. R., & Bayoumi, H. S. (2015). Bioremediation of different types of polluted water using microalgae. *Rendiconti Lincei*, *27*(2), 401–410. https://doi.org/10.1007/s12210-015-0495-1

Elahi, A., Arooj, I., Bukhari, D. A., & Rehman, A. (2020). Successive use of microorganisms to remove chromium from wastewater. *Applied Microbiology and Biotechnology*, *104*(9), 3729–3743. https://doi.org/10.1007/s00253-020-10533-y

Emparan, Q., Harun, R., & Danquah, M. K. (2019). Role of phycoremediation for nutrient removal from wastewaters: A review. *Applied Ecology and Environmental Research*, *17*(1), 889–915. https://doi.org/10.15666/aeer/1701_889915

Enamala, M. K., Enamala, S., Chavali, M., Donepudi, J., Yadavalli, R., Kolapalli, B., Aradhyula, T. V., Velpuri, J., & Kuppam, C. (2018). Production of biofuels from microalgae - A review on cultivation, harvesting, lipid extraction, and numerous applications of microalgae. *Renewable and Sustainable Energy Reviews*, *94*(May 2017), 49–68. https://doi.org/10.1016/j.rser.2018.05.012

Fernández, P. M., Viñarta, S. C., Bernal, A. R., Cruz, E. L., & Figueroa, L. I. C. (2018). Bioremediation strategies

Journal Pre-proof

for chromium removal: Current research, scale-up approach and future perspectives. In *Chemosphere* (Vol. 208, pp. 139–148). Elsevier Ltd. https://doi.org/10.1016/j.chemosphere.2018.05.166

Finnegan, P. M., & Chen, W. (2012). Arsenic toxicity: The effects on plant metabolism. *Frontiers in Physiology*, *3 JUN*(June), 1–18. https://doi.org/10.3389/fphys.2012.00182

Fomina, M., & Gadd, G. M. (2014). Biosorption: Current perspectives on concept, definition and application. *Bioresource Technology*, *160*, 3–14. https://doi.org/10.1016/j.biortech.2013.12.102

Forján, E., Navarro, F., Cuaresma, M., Vaquero, I., Ruíz-Domínguez, M. C., Gojkovic, Ž., Vázquez, M., Márquez, M., Mogedas, B., Bermejo, E., Girlich, S., Domínguez, M. J., Vílchez, C., Vega, J. M., & Garbayo, I. (2015). Microalgae: Fast-Growth Sustainable Green Factories. *Http://Dx.Doi.Org/10.1080/10643389.2014.966426*, *45*(16), 1705–1755. https://doi.org/10.1080/10643389.2014.966426

Gao, F., Cui, W., Xu, J. P., Li, C., Jin, W. H., & Yang, H. L. (2019). Lipid accumulation properties of Chlorella vulgaris and Scenedesmus obliquus in membrane photobioreactor (MPBR) fed with secondary effluent from municipal wastewater treatment plant. *Renewable Energy*, *136*, 671–676. https://doi.org/10.1016/j.renene.2019.01.038

Gao, Z., Jiang, C., Lyu, R., Yang, Z., & Zhang, T. (2020). Optimization of the preparation of fungal-algal pellets for use in the remediation of arsenic-contaminated water. *Environmental Science and Pollution Research*, *27*(29), 36789–36798. https://doi.org/10.1007/s11356-020-09757-2

Gaur, J. P., & Rai, L. C. (2001). Heavy Metal Tolerance in Algae. In *Algal Adaptation to Environmental Stresses* (pp. 363–388). https://doi.org/10.1007/978-3-642-59491-5_12

Genchi, G., Sinicropi, M. S., Lauria, G., Carocci, A., & Catalano, A. (2020). The effects of cadmium toxicity. *International Journal of Environmental Research and Public Health*, *17*(Cd), 1–24.

Guarín-Romero, J. R., Rodríguez-Estupiñán, P., Giraldo, L., & Moreno-Piraján, J. C. (2019). Simple and Competitive Adsorption Study of Nickel(II) and Chromium(III) on the Surface of the Brown Algae Durvillaea antarctica Biomass. *ACS Omega*, *4*(19), 18147–18158. https://doi.org/10.1021/acsomega.9b02061

Journal Pre-proof

Hamouda, R. A., El-Naggar, N. E. A., Doleib, N. M., & Saddiq, A. A. (2020). Bioprocessing strategies for cost-effective simultaneous removal of chromium and malachite green by marine alga Enteromorpha intestinalis. *Scientific Reports*, *10*(1), 1–19. https://doi.org/10.1038/s41598-020-70251-3

Han, T. W., Tseng, C. C., Cai, M., Chen, K., Cheng, S. Y., & Wang, J. (2020). Effects of cadmium on bioaccumulation, bioabsorption, and photosynthesis in Sarcodia suiae. *International Journal of Environmental Research and Public Health*, *17*(4), 1–12. https://doi.org/10.3390/ijerph17041294

Hassan, M. M., & Carr, C. M. (2018). A critical review on recent advancements of the removal of reactive dyes from dyehouse effluent by ion-exchange adsorbents. *Chemosphere*, *209*, 201–219. https://doi.org/10.1016/j.chemosphere.2018.06.043

Hedayatkhah, A., Cretoiu, M. S., Emtiazi, G., Stal, L. J., & Bolhuis, H. (2018). Bioremediation of chromium contaminated water by diatoms with concomitant lipid accumulation for biofuel production. *Journal of Environmental Management*, *227*(July), 313–320. https://doi.org/10.1016/j.jenvman.2018.09.011

Henriques, B., Teixeira, A., Figueira, P., Reis, A. T., Almeida, J., Vale, C., & Pereira, E. (2019). Simultaneous removal of trace elements from contaminated waters by living Ulva lactuca. *Science of the Total Environment*, *652*, 880–888. https://doi.org/10.1016/j.scitotenv.2018.10.282

Huang, R., Huo, G., Song, S., Li, Y., Xia, L., & Gaillard, J. F. (2019). Immobilization of mercury using high-phosphate culture-modified microalgae. *Environmental Pollution*, *254*, 1–10. https://doi.org/10.1016/j.envpol.2019.112966

Hussain, M. M., Wang, J., Bibi, I., Shahid, M., Niazi, N. K., Iqbal, J., Mian, I. A., Shaheen, S. M., Bashir, S., Shah, N. S., Hina, K., & Rinklebe, J. (2021). Arsenic speciation and biotransformation pathways in the aquatic ecosystem: The significance of algae. *Journal of Hazardous Materials*, *403*, 1–59. https://doi.org/10.1016/j.jhazmat.2020.124027

Jais, N. M., Mohamed, R. M. S. R., Al-Gheethi, A. A., & Hashim, M. K. A. (2017). The dual roles of phycoremediation of wet market wastewater for nutrients and heavy metals removal and microalgae biomass production. *Clean Technologies and Environmental Policy*, *19*(1), 37–52. https://doi.org/10.1007/s10098-016-

Journal Pre-proof

1235-7

Jie, W., Zhenyu, L., Xin, H., Tao, Y., Mengliang, L., Qiubai, D., Khalid, K., Zhongyuan, L., & Shuzhen, L. (2021). Porous Materials Prepared by Magnesium Phosphate Cement for the Effective Immobilization of Lead Ions. *International Journal of Environmental Research*, 1–14. https://doi.org/10.1007/s41742-021-00344-y

Joutey, N. T., Sayel, H., Bahafid, W., & Ghachtouli, E. (2014). Mechanisms of Hexavalent Chromium Resistance and Removal by Microorganisms. *Reviews of Environmental Contamination and Toxicology*, *233*, 45–69. https://doi.org/10.1007/978-3-319-10479-9

Kaloudas, D., Pavlova, N., & Penchovsky, R. (2021). Phycoremediation of wastewater by microalgae: a review. *Environmental Chemistry Letters*, *19*(4), 2905–2920. https://doi.org/10.1007/S10311-021-01203-0/TABLES/2

Kapahi, M., & Sachdeva, S. (2019). Bioremediation options for heavy metal pollution. *Journal of Health and Pollution*, *9*(24), 1–20. https://doi.org/10.5696/2156-9614-9.24.191203

Kesari, K. K., Soni, R., Jamal, Q. M. S., Tripathi, P., Lal, J. A., Jha, N. K., Siddiqui, M. H., Kumar, P., Tripathi, V., & Ruokolainen, J. (2021). Wastewater Treatment and Reuse: a Review of its Applications and Health Implications. *Water, Air, & Soil Pollution*, *232*(5), 208. https://doi.org/10.1007/s11270-021-05154-8

Khajavian, M., Wood, D. A., Hallajsani, A., & Majidian, N. (2019). Simultaneous biosorption of nickel and cadmium by the brown algae Cystoseria indica characterized by isotherm and kinetic models. *Applied Biological Chemistry*, *62*(1). https://doi.org/10.1186/s13765-019-0477-6

Khalid, S., Shahid, M., Natasha, Bibi, I., Sarwar, T., Shah, A. H., & Niazi, N. K. (2018). A review of environmental contamination and health risk assessment of wastewater use for crop irrigation with a focus on low and high-income countries. *International Journal of Environmental Research and Public Health*, *15*(5), 1–36. https://doi.org/10.3390/ijerph15050895

Kikuchi, T., & Tanaka, S. (2012). Biological Removal and Recovery of Toxic Heavy Metals in Water Environment. *Http://Dx.Doi.Org/10.1080/10643389.2011.651343*, *42*(10), 1007–1057. https://doi.org/10.1080/10643389.2011.651343

Journal Pre-proof

Koreiviene, J., Karosiene, J., Kasperoviciene, J., Paskauskas, R., Leska, B., Pankiewicz, R., Juskaite, L., Zagorskis, A., Wilk-woz-niak, E., Valskys, V., Gulbinas, Z., Walusiak, E., Krzton, W., Morudov, D., Radzevicius, K., Treska, E., Tabisz, L., Papsdorf, M., Piotrowicz, Z., & Messyasz, B. (2019). Algae service for life develops ecologically sustainable bioproducts from freshwater cyanobacteria and macroalgae biomass. *EU Project Of Life Programme*, *25*(2), 176–185.

Kumar, L., & Bharadvaja, N. (2020a). A review on microalgae biofuel and biorefinery: challenges and way forward. *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, *00*(00), 1–24. https://doi.org/10.1080/15567036.2020.1836084

Kumar, L., & Bharadvaja, N. (2020b). Microbial Remediation of Heavy Metals. In *Microbial Bioremediation & Biodegradation* (pp. 49–72). Springer Singapore. https://doi.org/10.1007/978-981-15-1812-6_2

Kumar, L., & Bharadvaja, N. (2021). Algal-Based Wastewater Treatment and Biorefinery. *Wastewater Treatment*, 413–432. https://doi.org/10.1016/b978-0-12-821881-5.00020-9

Kumar, L., Chugh, M., Kumar, S., Kumar, K., Sharma, J., & Bharadvaja, N. (2022). Remediation of petrorefinery wastewater contaminants: A review on physicochemical and bioremediation strategies. *Process Safety and Environmental Protection*, *0957–5820*. https://doi.org/10.1016/J.PSEP.2022.01.009

Kumar, L., Khushbu, Chugh, M., & Bharadvaja, N. (2022). Microbial remediation of tannery wastewater. *Development in Wastewater Treatment Research and Processes*, 303–328. https://doi.org/10.1016/B978-0-323-85657-7.00011-0

Lavrinovičs, A., & Juhna, T. (2018). Review on Challenges and Limitations for Algae-Based Wastewater Treatment. *Construction Science*, *20*(1), 17–25. https://doi.org/10.2478/cons-2017-0003

Lebrun, M., Miard, F., Nandillon, R., Morabito, D., & Bourgerie, S. (2021). Biochar Application Rate: Improving Soil Fertility and Linum usitatissimum Growth on an Arsenic and Lead Contaminated Technosol. *International Journal of Environmental Research*, *15*(1), 125–134. https://doi.org/10.1007/s41742-020-00302-0

Lee, C. S., Oh, H. S., Oh, H. M., Kim, H. S., & Ahn, C. Y. (2016). Two-phase photoperiodic cultivation of algal-

bacterial consortia for high biomass production and efficient nutrient removal from municipal wastewater. *Bioresource Technology*, *200*, 867–875. https://doi.org/10.1016/j.biortech.2015.11.007

Lee, L., Hsu, C. Y., & Yen, H. W. (2017). The effects of hydraulic retention time (HRT) on chromium(VI) reduction using autotrophic cultivation of Chlorella vulgaris. *Bioprocess and Biosystems Engineering*, *40*(12), 1725–1731. https://doi.org/10.1007/s00449-017-1827-6

Leong, Y. K., & Chang, J. S. (2020). Bioremediation of heavy metals using microalgae: Recent advances and mechanisms. *Bioresource Technology*, *303*(January), 1–11. https://doi.org/10.1016/j.biortech.2020.122886

Libatique, M. J. H., Lee, M. C., Yeh, H. Y., & Jhang, F. J. (2020). Total and inorganic arsenic biosorption by Sarcodia suiae (Rhodophyta), as affected by controlled environmental conditions. *Chemosphere*, *248*, 1–10. https://doi.org/10.1016/j.chemosphere.2020.126084

Logroño, W., Pérez, M., Urquizo, G., Kadier, A., Echeverría, M., Recalde, C., & Rákhely, G. (2017). Single chamber microbial fuel cell (SCMFC) with a cathodic microalgal biofilm: A preliminary assessment of the generation of bioelectricity and biodegradation of real dye textile wastewater. *Chemosphere*, *176*, 378–388. https://doi.org/10.1016/j.chemosphere.2017.02.099

Luo, Y., Le-Clech, P., & Henderson, R. K. (2017). Simultaneous microalgae cultivation and wastewater treatment in submerged membrane photobioreactors: A review. *Algal Research*, *24*, 425–437. https://doi.org/10.1016/j.algal.2016.10.026

Mansilla, A., Ávila, M., & Yokoya, N. S. (2012). Current knowledge on biotechnological interesting seaweeds from the Magellan Region, Chile. *Brazilian Journal of Pharmacognosy*, *22*(4), 760–767. https://doi.org/10.1590/S0102-695X2012005000074

Massocato, T. F., Ramos, J. C., Bascuñan, V. L. F., Simioni, C., Rörig, L. R., & Bonomi Barufi, J. (2018). Tolerance of Ulothrix sp. LAFIC 010 (Chlorophyta) against high concentration of metals from acid mine drainage. *Ecotoxicology and Environmental Safety*, *157*(October 2017), 227–234. https://doi.org/10.1016/j.ecoenv.2018.03.045

Mohammady, N. G.-E., El-Sayed, H. S., Fakhry, E. M., Taha, H. M., Mohamed, J. H., Mahmoud, N. H., &

Abdelsalam, B. A. (2015). Locally Isolated Microalgae as a Source of Biodiesel and By- Products: An Integral Study of Med-Algae Project. *International Journal of Chemical Concepts*, *1*(2395–4256), 94–102. http://www.iglobaljournal.com/wp-content/uploads/2014/08/4.-Bhattacharjee-et-al.-IGJPS-2014.pdf%5Cnhttp://dx.doi.org/10.1007/s13410-013-0125-3%5Cnhttp://dx.doi.org/10.1007/s12562-012-0552-y%5Cnhttp://dx.doi.org/10.1007/BF02976957%5Cnhttp://dx.doi.org/10.1

Mokone, J. G., Tutu, H., Chimuka, L., & Cukrowska, E. M. (2018). Optimization and Characterization of Cladophora sp. Alga Immobilized in Alginate Beads and Silica Gel for the Biosorption of Mercury from Aqueous Solutions. *Water, Air, and Soil Pollution*, *229*(7), 1–14. https://doi.org/10.1007/s11270-018-3859-1

Molino, A., Mehariya, S., Karatza, D., Chianese, S., Iovine, A., Casella, P., Marino, T., & Musmarra, D. (2019). Bench-scale cultivation of microalgae scenedesmus almeriensis for CO2 capture and lutein production. *Energies*, *12*(14), 1–14. https://doi.org/10.3390/en12142806

Moreira, V. R., Lebron, Y. A. ., Lange, L., & Santos, L. V. S. (2019). Simultaneous biosorption of Cd(II), Ni(II) and Pb(II) onto a brown macroalgae Fucus vesiculosus: Mono- and multi-component isotherms, kinetics and thermodynamics. *Journal of Environmental Management*, *251*(May), 109587. https://doi.org/10.1016/j.jenvman.2019.109587

Nájera-Martínez, M., Pérez-Cruz, A., Dzul-Caamal, R., & Vega-López, A. (2021). Are the Endogenous Levels of Divalent Heavy Metals Responsible for the Oxidative Stress Response on Freshwater Phytoplankton Communities? *Water, Air, and Soil Pollution*, *232*(2), 1–12. https://doi.org/10.1007/s11270-021-05035-0

Navakoudis, A. M. E., & Ververidis, K. P. F. (2018). Microalgae : a potential tool for remediating aquatic environments from toxic metals. *International Journal of Environmental Science and Technology*, *0123456789*. https://doi.org/10.1007/s13762-018-1783-y

Naveed, S., Li, C., Lu, X., Chen, S., Yin, B., Zhang, C., & Ge, Y. (2019). Microalgal extracellular polymeric substances and their interactions with metal(loid)s: A review. *Https://Doi.Org/10.1080/10643389.2019.1583052*, *49*(19), 1769–1802. https://doi.org/10.1080/10643389.2019.1583052

Nguyen, H. T., Yoon, Y., Ngo, H. H., & Jang, A. (2020). The application of microalgae in removing organic micropollutants in wastewater. *Https://Doi.Org/10.1080/10643389.2020.1753633*, *51*(12), 1187–1220. https://doi.org/10.1080/10643389.2020.1753633

Nowicka, B., Fesenko, T., Walczak, J., & Kruk, J. (2020). The inhibitor-evoked shortage of tocopherol and plastoquinol is compensated by other antioxidant mechanisms in Chlamydomonas reinhardtii exposed to toxic concentrations of cadmium and chromium ions. *Ecotoxicology and Environmental Safety*, *191*(October 2019), 1–12. https://doi.org/10.1016/j.ecoenv.2020.110241

Nur, M. M. A., & Buma, A. G. J. (2019). Opportunities and Challenges of Microalgal Cultivation on Wastewater, with Special Focus on Palm Oil Mill Effluent and the Production of High Value Compounds. *Waste and Biomass Valorization*, *10*(8), 2079–2097. https://doi.org/10.1007/s12649-018-0256-3

Nzayisenga, J. C., Farge, X., Groll, S. L., & Sellstedt, A. (2020). Effects of light intensity on growth and lipid production in microalgae grown in wastewater. *Biotechnology for Biofuels*, *13*(1), 1–9. https://doi.org/10.1186/s13068-019-1646-x

Pacheco, D., Rocha, A. C., Pereira, L., & Verdelhos, T. (2020). Microalgae water bioremediation: Trends and hot topics. *Applied Sciences (Switzerland)*, *10*(5). https://doi.org/10.3390/app10051886

Park, J. H., Hong, J. Y., Jang, H. C., Oh, S. G., Kim, S. H., Yoon, J. J., & Kim, Y. J. (2012). Use of Gelidium amansii as a promising resource for bioethanol: A practical approach for continuous dilute-acid hydrolysis and fermentation. *Bioresource Technology*, *108*, 83–88. https://doi.org/10.1016/j.biortech.2011.12.065

Patidar, J., Sharma, R., Yadav, M., & Tiwari, A. (2017). Microalgae as sustainable renewable energy feedstock for bioethanol production. *Scholars Academic Journal of Biosciences*, *5*(6), 459–464. https://doi.org/10.21276/sajb

Patra, M., & Sharma, A. (2000). Mercury toxicity in plants. *Botanical Review*, *66*(3), 379–422. https://doi.org/10.1007/BF02868923

Peng, Y., Deng, A., Gong, X., Li, X., & Zhang, Y. (2017). Coupling process study of lipid production and mercury bioremediation by biomimetic mineralized microalgae. *Bioresource Technology*, *243*, 628–633.

https://doi.org/10.1016/j.biortech.2017.06.165

Perales-Vela, H. V., Pen˜a-Castro, J. M., & Can˜izares-Villanueva, R. O. (2006). Heavy metal detoxification in eukaryotic microalgae. *Chemosphere*, *64*, 1–10. https://doi.org/10.1016/j.chemosphere.2005.11.024

Plácido, J., Bustamante-López, S., Meissner, K. E., Kelly, D. E., & Kelly, S. L. (2019). Microalgae biochar-derived carbon dots and their application in heavy metal sensing in aqueous systems. *Science of the Total Environment*, *656*, 531–539. https://doi.org/10.1016/j.scitotenv.2018.11.393

Pratush, A. (2018). Adverse effect of heavy metals (As , Pb , Hg , and Cr) on health and their bioremediation strategies : a review. *International Microbiology*.

Priyadarshani, I., Sahu, D., & Rath, B. (2012). Microalgal bioremediation : Current practices and perspectives. *Journal of Biochemical Technology*, *3*(3), 299–304.

Priyadarshini, E., Priyadarshini, S. S., & Pradhan, N. (2019). Heavy metal resistance in algae and its application for metal nanoparticle synthesis. *Applied Microbiology and Biotechnology*, *103*(8), 3297–3316. https://doi.org/10.1007/s00253-019-09685-3

Qian, J., Wei, L., Liu, R., Jiang, F., Hao, X., & Chen, G. H. (2016). An exploratory study on the pathways of Cr (VI) reduction in sulfate-reducing Up-flow Anaerobic Sludge Bed (UASB) reactor. *Scientific Reports*, *6*(March), 1–12. https://doi.org/10.1038/srep23694

Quijano, G., Arcila, J. S., & Buitrón, G. (2017). Microalgal-bacterial aggregates: Applications and perspectives for wastewater treatment. *Biotechnology Advances*, *35*(6), 772–781. https://doi.org/10.1016/j.biotechadv.2017.07.003

Radwan, E. K., Abdel-Aty, A. M., El-Wakeel, S. T., & Abdel Ghafar, H. H. (2020). Bioremediation of potentially toxic metal and reactive dye-contaminated water by pristine and modified Chlorella vulgaris. *Environmental Science and Pollution Research*, *27*(17), 21777–21789. https://doi.org/10.1007/s11356-020-08550-5

Rajamohan, R., Rao, T. S., Anupkumar, B., Sahayam, A. C., Krishna, M. V. B., Venugopalan, V. P., & Narasimhan, S. V. (2010). Distribution of heavy metals in the vicinity of a nuclear power plant, east coast of India: With

Journal Pre-proof

emphasis on copper concentration and primary productivity. *Indian Journal of Marine Sciences*, *39*(2), 182–191.

Ramakrishnan, B., Megharaj, M., Venkateswarlu, K., Naidu, R., & Sethunathan, N. (2010). The Impacts of Environmental Pollutants on Microalgae and Cyanobacteria. *Http://Dx.Doi.Org/10.1080/10643380802471068*, *40*(8), 699–821. https://doi.org/10.1080/10643380802471068

Ramírez-Rodríguez, A. E., Bañuelos-Hernández, B., García-Soto, M. J., Govea-Alonso, D. G., Rosales-Mendoza, S., Alfaro de la Torre, M. C., Monreal-Escalante, E., & Paz-Maldonado, L. M. T. (2019). Arsenic removal using Chlamydomonas reinhardtii modified with the gene acr3 and enhancement of its performance by decreasing phosphate in the growing media. *International Journal of Phytoremediation*, *21*(7), 617–623. https://doi.org/10.1080/15226514.2018.1546274

Ravindran, B., Gupta, S. K., Cho, W. M., Kim, J. K., Lee, S. R., Jeong, K. H., Lee, D. J., & Choi, H. C. (2016). Microalgae potential and multiple roles-current progress and future prospects-an overview. *Sustainability (Switzerland)*, *8*(12), 1–16. https://doi.org/10.3390/su8121215

Rezvani, F., Sarrafzadeh, M. H., Seo, S. H., & Oh, H. M. (2018). Optimal strategies for bioremediation of nitrate-contaminated groundwater and microalgae biomass production. *Environmental Science and Pollution Research*, *25*(27), 27471–27482. https://doi.org/10.1007/s11356-018-2777-5

Saavedra, R., Muñoz, R., Taboada, M. E., Vega, M., & Bolado, S. (2018). Comparative uptake study of arsenic, boron, copper, manganese and zinc from water by different green microalgae. *Bioresource Technology*, *263*, 49–57. https://doi.org/10.1016/j.biortech.2018.04.101

Salam, L. B., Shomope, H., Ummi, Z., & Bukar, F. (2019). Mercury contamination imposes structural shift on the microbial community of an agricultural soil. *Bulletin of the National Research Centre*, *43*(1), 1–13. https://doi.org/10.1186/s42269-019-0208-5

Salehi, B., Sharifi-rad, J., Seca, A. M. L., & Pinto, D. C. G. A. (2019). Current Trends on Seaweeds : Looking at Chemical composition, phytopharmacology, and cosmetic applications. *Molecules*, *24*(22), 4182.

Journal Pre-proof

Saranya, D., & Shanthakumar, S. (2019). Green microalgae for combined sewage and tannery effluent treatment: Performance and lipid accumulation potential. *Journal of Environmental Management*, *241*(April), 167–178. https://doi.org/10.1016/j.jenvman.2019.04.031

Sayed, E., Hyun, S., Roh, S., Dev, S., Ali, M., Reda, K., & Shanab, A. I. A. (2019). Algae as a green technology for heavy metals removal from various wastewater. *World Journal of Microbiology and Biotechnology*, *35*(5), 1–19. https://doi.org/10.1007/s11274-019-2648-3

Scranton, M. A., Ostrand, J. T., Fields, F. J., & Mayfield, S. P. (2017a). Chlamydomonas as a model for biofuels and bio-products production. *Plant J*, *82*(3), 523–531. https://doi.org/10.1111/tpj.12780.Chlamydomonas

Scranton, M. A., Ostrand, J. T., Fields, F. J., & Mayfield, S. P. (2017b). *Chlamydomonas as a model for biofuels and bio-products production*. *82*(3), 523–531. https://doi.org/10.1111/tpj.12780.Chlamydomonas

Senthilkumar, R., Reddy Prasad, D. M., Govindarajan, L., Saravanakumar, K., & Naveen Prasad, B. S. (2020). Synthesis of green marine algal-based biochar for remediation of arsenic(V) from contaminated waters in batch and column mode of operation. *International Journal of Phytoremediation*, *22*(3), 279–286. https://doi.org/10.1080/15226514.2019.1658710

Shamim, S. (2018). Biosorption of Heavy Metals. In *Biosorption, intech* (Issue July, pp. 21–49). http://www.intechopen.com/books/trends-in-telecommunications-technologies/gps-total-electron-content-tec-prediction-at-ionosphere-layer-over-the-equatorial-region%0AInTec%0Ahttp://www.asociatiamhc.ro/wp-content/uploads/2013/11/Guide-to-Hydropower.pdf

Shamshad, I., Khan, S., Waqas, M., Asma, M., Nawab, J., Gul, N., Raiz, A., & Li, G. (2016). Heavy metal uptake capacity of fresh water algae (Oedogonium westti) from aqueous solution: A mesocosm research. *International Journal of Phytoremediation*, *18*(4), 393–398. https://doi.org/10.1080/15226514.2015.1109594

Shanab, S., Essa, A., & Shalaby, E. (2012). Bioremoval capacity of three heavy metals by some microalgae species (Egyptian isolates). *Plant Signaling and Behavior*, *7*(3). https://doi.org/10.4161/psb.19173

Shi, W., Jin, Z., Hu, S., Fang, X., & Li, F. (2017). Dissolved organic matter affects the bioaccumulation of copper and lead in Chlorella pyrenoidosa: A case of long-term exposure. *Chemosphere*, *174*, 447–455.

https://doi.org/10.1016/j.chemosphere.2017.01.119

Siddeeg, S. M., Tahoon, M. A., & Ben Rebah, F. (2020). Agro-industrial waste materials and wastewater as growth media for microbial bioflocculants production: A review. *Materials Research Express*, *7*(1). https://doi.org/10.1088/2053-1591/ab5980

Singh, A., Kumar, D., & Gaur, J. P. (2012). Continuous metal removal from solution and industrial effluents using Spirogyra biomass-packed column reactor. *Water Research*, *46*(3), 779–788. https://doi.org/10.1016/j.watres.2011.11.050

Singh, D. V., Bhat, R. A., Upadhyay, A. K., Singh, R., & Singh, D. P. (2021). Microalgae in aquatic environs: A sustainable approach for remediation of heavy metals and emerging contaminants. *Environmental Technology & Innovation*, *21*, 101340. https://doi.org/10.1016/J.ETI.2020.101340

Solisio, C., Al Arni, S., & Converti, A. (2019). Adsorption of inorganic mercury from aqueous solutions onto dry biomass of Chlorella vulgaris: kinetic and isotherm study. *Environmental Technology (United Kingdom)*, *40*(5), 664–672. https://doi.org/10.1080/09593330.2017.1400114

Tabassum, M. R., Xia, A., & Murphy, J. D. (2017). Potential of seaweed as a feedstock for renewable gaseous fuel production in Ireland. *Renewable and Sustainable Energy Reviews*, *68*(September 2016), 136–146. https://doi.org/10.1016/j.rser.2016.09.111

Tu, C., Liu, Y., Wei, J., Li, L., Scheckel, K. G., & Luo, Y. (2018). Characterization and mechanism of copper biosorption by a highly copper-resistant fungal strain isolated from copper-polluted acidic orchard soil. *Environmental Science and Pollution Research*, *25*(25), 24965–24974. https://doi.org/10.1007/s11356-018-2563-4

Tumolo, M., Ancona, V., De Paola, D., Losacco, D., Campanale, C., Massarelli, C., & Uricchio, V. F. (2020). Chromium pollution in European water, sources, health risk, and remediation strategies: An overview. *International Journal of Environmental Research and Public Health*, *17*(15), 1–25. https://doi.org/10.3390/ijerph17155438

Turner, A., Brice, D., & Brown, M. T. (2012). Interactions of silver nanoparticles with the marine macroalga, Ulva

Journal Pre-proof

lactuca. *Ecotoxicology*, *21*(1), 148–154. https://doi.org/10.1007/s10646-011-0774-2

Ugya, A. Y., Ajibade, F. O., & Hua, X. (2021). The efficiency of microalgae biofilm in the phycoremediation of water from River Kaduna. *Journal of Environmental Management*, *295*. https://doi.org/10.1016/J.JENVMAN.2021.113109

UNESCO. (2015). *The United Nations World Water Development Report 2015: Water for a Sustainable World. Paris, UNESCO*.

Upadhyay, A. K., Mandotra, S. K., Kumar, N., Singh, N. K., Singh, L., & Rai, U. N. (2016). Augmentation of arsenic enhances lipid yield and defense responses in alga Nannochloropsis sp. *Bioresource Technology*, *221*(2016), 430–437. https://doi.org/10.1016/j.biortech.2016.09.061

Upadhyay, Atul Kumar, Singh, L., Singh, R., Singh, D. P., & Saxena, G. (2020). Bioaccumulation and Toxicity of As in the Alga Chlorococcum sp.: Prospects for As Bioremediation. *Bulletin of Environmental Contamination and Toxicology*, *0123456789*, 1–7. https://doi.org/10.1007/s00128-020-02964-0

Usmani, Z., Sharma, M., Lukk, T., Karpichev, Y., Thakur, V. K., Kumar, V., Allaoui, A., Awasthi, A. K., & Gupta, V. K. (2020). Developments in enzyme and microalgae based biotechniques to remediate micropollutants from aqueous systems—A review. *Https://Doi.Org/10.1080/10643389.2020.1862551*. https://doi.org/10.1080/10643389.2020.1862551

Utomo, H. D., Tan, K. X. D., Choong, Z. Y. D., Yu, J. J., Ong, J. J., & Lim, Z. B. (2016). Biosorption of Heavy Metal by Algae Biomass in Surface Water. *Journal of Environmental Protection*, *07*(11), 1547–1560. https://doi.org/10.4236/jep.2016.711128

Vassilev, S. V., & Vassileva, C. G. (2016). Composition, properties and challenges of algae biomass for biofuel application: An overview. *Fuel*, *181*, 1–33. https://doi.org/10.1016/j.fuel.2016.04.106

Vo, H. N. P., Ngo, H. H., Guo, W., Chang, S. W., Nguyen, D. D., Chen, Z., Wang, X. C., Chen, R., & Zhang, X. (2019). Microalgae for saline wastewater treatment: a critical review. *Https://Doi.Org/10.1080/10643389.2019.1656510*, *50*(12), 1224–1265. https://doi.org/10.1080/10643389.2019.1656510

Journal Pre-proof

Vo, M.-T., Nguyen, V.-T., Vo, T.-M.-C., Bui, T.-N.-P., & Dao, T.-S. (2020). Responses of green algae and diatom upon exposure to chromium and cadmium. *Vietnam Journal of Science, Technology and Engineering*, *62*(1), 69–73. https://doi.org/10.31276/vjste.62(1).69-73

Wei, S., Cao, J., Ma, X., Ping, J., Zhang, C., Ke, T., Zhang, Y., Tao, Y., & Chen, L. (2020). The simultaneous removal of the combined pollutants of hexavalent chromium and o-nitrophenol by Chlamydomonas reinhardtii. *Ecotoxicology and Environmental Safety*, *198*(May), 1–9. https://doi.org/10.1016/j.ecoenv.2020.110648

Wu, J., Ma, L. L., & Zeng, R. J. (2018). Role of extracellular polymeric substances in efficient chromium(VI) removal by algae-based Fe/C nano-composite. *Chemosphere*, *211*, 608–616. https://doi.org/10.1016/j.chemosphere.2018.07.186

Wu, Z. Z., Li, D. Y., & Cheng, Y. S. (2018). Application of ensilage as a green approach for simultaneous preservation and pretreatment of macroalgae Ulva lactuca for fermentable sugar production. *Clean Technologies and Environmental Policy*, *20*(9), 2057–2065. https://doi.org/10.1007/s10098-018-1574-7

Xie, Y., Ji, X., Wu, J., Tian, F., Zhu, J., & Liu, Z. (2021). Assessment of Cadmium (Cd) in Paddy Soil and Ditch Sediment in Polluted Watershed and Non-Polluted Watershed. *International Journal of Environmental Research*, *15*(3), 527–534. https://doi.org/10.1007/s41742-021-00327-z

Yalçın, S., & Özyürek, M. (2018). Biosorption potential of two brown seaweeds in the removal of chromium. *Water Science and Technology*, *78*(12), 2564–2576. https://doi.org/10.2166/wst.2019.007

Yang, J. S., Cao, J., Xing, G. L., & Yuan, H. L. (2015). Lipid production combined with biosorption and bioaccumulation of cadmium, copper, manganese and zinc by oleaginous microalgae Chlorella minutissima UTEX2341. *Bioresource Technology*, *175*, 537–544. https://doi.org/10.1016/j.biortech.2014.10.124

Yasin, G., Ur Rahman, S., Yousaf, M. T. Bin, Azhar, M. F., Zahid, D. M., Imtiaz, M., & Hussain, B. (2021). Phytoremediation Potential of E. camaldulensis and M. alba for Copper, Cadmium, and Lead Absorption in Urban Areas of Faisalabad City, Pakistan. *International Journal of Environmental Research*, 1–16. https://doi.org/10.1007/s41742-021-00330-4

Journal Pre-proof

Ye, J., Xiao, H., Xiao, B., Xu, W., Gao, L., & Lin, G. (2015). Bioremediation of heavy metal contaminated aqueous solution by using red algae Porphyra leucosticta. *Water Science and Technology*, *72*(9), 1662–1666. https://doi.org/10.2166/wst.2015.386

Yen, H. W., Chen, P. W., Hsu, C. Y., & Lee, L. (2017). The use of autotrophic Chlorella vulgaris in chromium (VI) reduction under different reduction conditions. *Journal of the Taiwan Institute of Chemical Engineers*, *74*, 1–6. https://doi.org/10.1016/j.jtice.2016.08.017

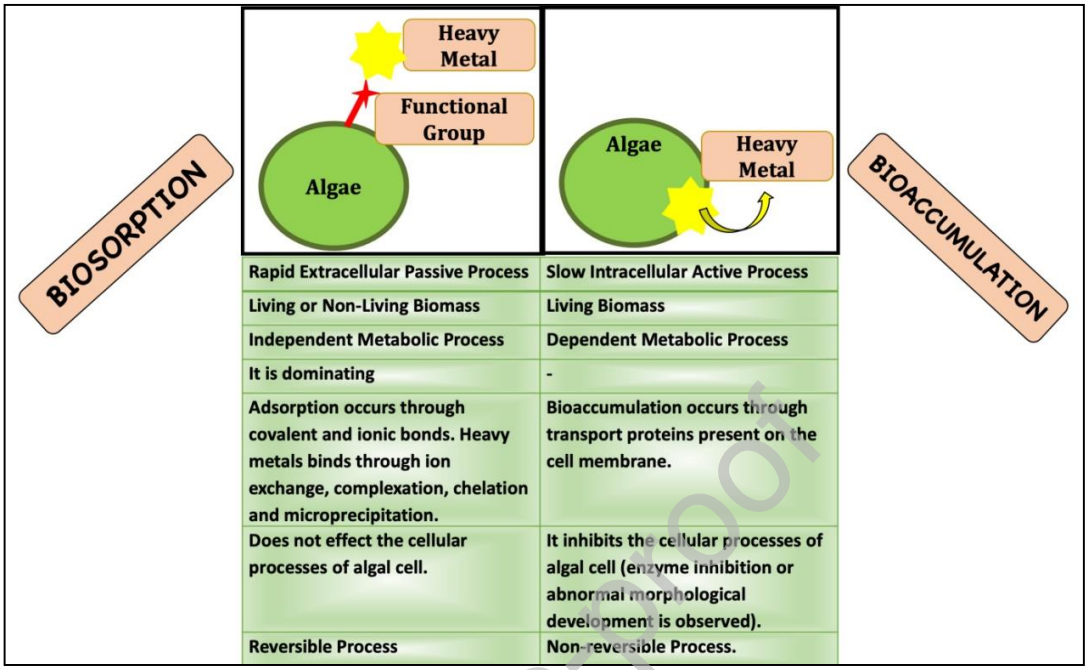
Zhang, J., Zhou, F., Liu, Y., Huang, F., & Zhang, C. (2019). Effect of extracellular polymeric substances on arsenic accumulation in Chlorella pyrenoidosa. *Science of the Total Environment*, *704*(xxxx), 135368. https://doi.org/10.1016/j.scitotenv.2019.135368

Zhao, Y., Song, X., Zhong, D. bo, Yu, L., & Yu, X. (2019). γ-Aminobutyric acid (GABA) regulates lipid production and cadmium uptake by Monoraphidium sp. QLY-1 under cadmium stress. *Bioresource Technology*, *297*. https://doi.org/10.1016/j.biortech.2019.122500

Zohoorian, H., Ahmadzadeh, H., Molazadeh, M., Shourian, M., & Lyon, S. (2020). Microalgal bioremediation of heavy metals and dyes. In *Algal Bioremediation* (pp. 659–674). Elsevier Inc. https://doi.org/10.1016/b978-0-12-818305-2.00041-3

Journal Pre-proof

Figures



Rapid Extracellular Passive Process	Slow Intracellular Active Process
Living or Non-Living Biomass	Living Biomass
Independent Metabolic Process	Dependent Metabolic Process
It is dominating	-
Adsorption occurs through covalent and ionic bonds. Heavy metals binds through ion exchange, complexation, chelation and microprecipitation.	Bioaccumulation occurs through transport proteins present on the cell membrane.
Does not effect the cellular processes of algal cell.	It inhibits the cellular processes of algal cell (enzyme inhibition or abnormal morphological development is observed).
Reversible Process	Non-reversible Process.

Fig.1 Two stage mechanism: remediation of heavy metals

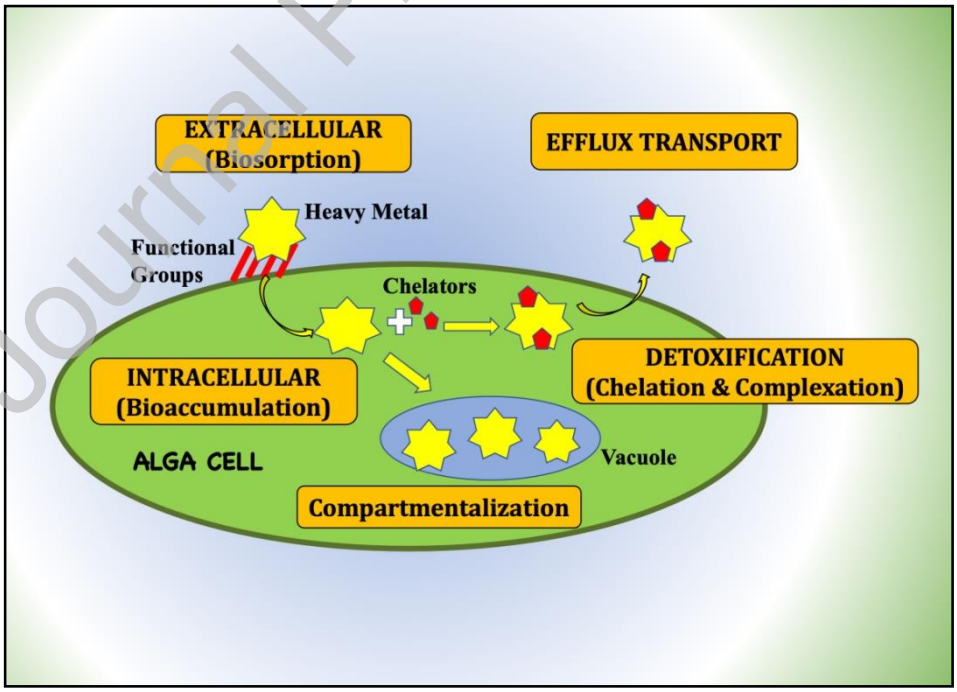


Fig. 2 Self-defense mechanisms of algae against heavy metal toxicity

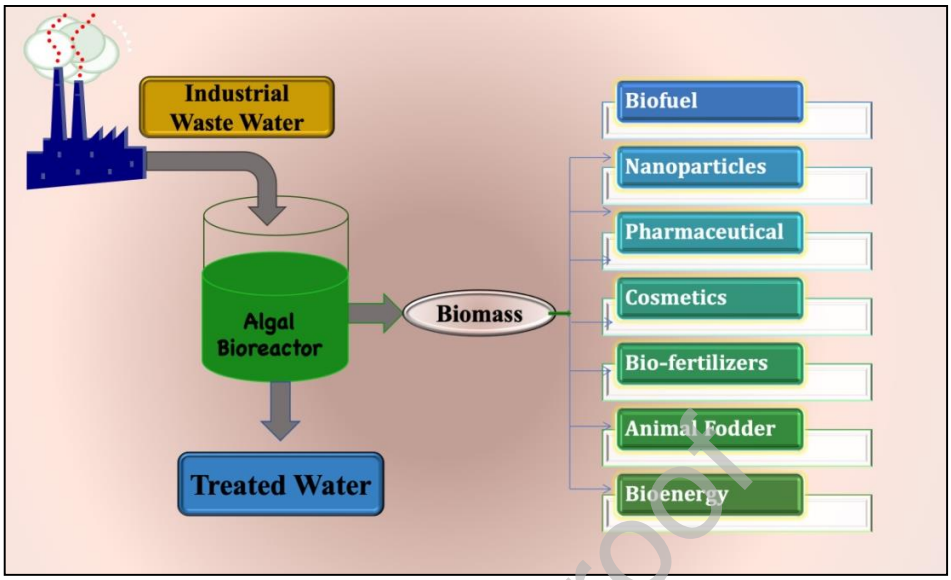


Fig. 3 Economical set-up for the remediation of industrial wastewater

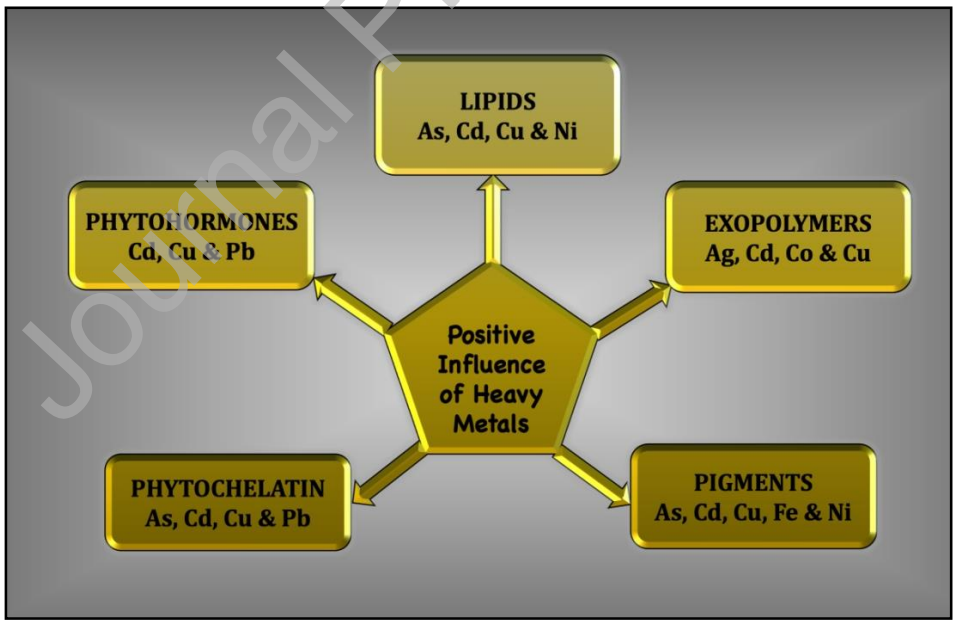


Fig. 4 Heavy metals enhancing the accumulation of by-products

Journal Pre-proof

Table 1 Phycoremediation of heavy metals along with the removal efficiency, time, mechanism, and growth parameters

Type of pollutants	Algae strain	Growth conditions	Remediation Performance	Bioremediation Time	Mechanism	Other remarks	Reference
Arsenic	Scenedesmus almeriensis	23 ± 2ºC, 9.5 pH, 12L:12D (1000 µE/m^2/s), 250 rpm.	40.7%	3 h	-	Removal Efficiency increases in multimetallic solution.	(Saavedra et al., 2018)
	Ulva lactuca	20 ± 2 °C, 7.8 pH, 12L:12D (natural Light.	48%	12 h	Biosorption and bioaccumulation	No effect of salinity.	(Henriques et al., 2019)
	Chlamydomonas reinhardtii acr3 modified strain	25°C, 16L: 8D, 150 rpm	1.5 – 3 times the wild strain.	-	-	-	(Ramírez-Rodríguez et al., 2019)
	Sarcodia suiae	15, 80 mmol photons m^{-2} s^{-1} (12L:12D)	-	-	Biosorption	Light have no significant effect.	(Libatique et al., 2020)
	Ulva Reticulata (Biochar)	30±1 °C, 4-6 pH, 30 mins, 180 rpm	8.12mg/g or 59.5%	30 mins	Biosorption	Desorption 99.5%	(Senthilkumar et al., 2020)
Boron	Scenedesmus almeriensisis	23 ± 2ºC, 5.5 pH, 12L:12D (1000 µE/m^2/s) and 250 rpm.	38.6%	10 mins	-	Growth was effect.	(Saavedra et al., 2018)
Cadmium	*Phormidium ambiguum*	20 ± 1°C, 30 mE/m^2 /s, (16L:8D), 130 M/min	86%	-	Biosorption and Bioaccumulation	Inhibitory effects towards growth.	(Shanab et al., 2012)
	Desmodesmus sp. MAS1 and Hetero	23 ± 1 °C, 3.5 pH, 60 µmol photons m^{-2} s^{-1}, 100 rpm.	More than 58%	16 days	Biosorption and Bioaccumulation.	Produce Biodiesel.	(Abinandan et al., 2019)

Journal Pre-proof

	chlorella sp. MAS3						
	Chlorococcum humicola	25 ± 2 °C, 2000–3000 lux (14L: 10 D).	17%	6 days	Biosorption	binds with the hydroxyl, thiol or carboxyl group.	(Borah et al., 2020)
	Porphyra leucosticta	30 °C and 150 rpm.	synthetic waste water (cd) : 70% real industrial waste water (Cd): 75%	2 h	-	Bioremediation increases with increase in contact time.	(Ye et al., 2015)
	Oedogonium westi	18±1°C, 5 pH, natural light (12:12h).	95%	7 days	Biosorptiom	-	(Shamshad et al., 2016)
Chromium	*Durvillaea antarctica*	20 ± 1°C, 5 pH, 160 rpm.	102.72 mg/g	420 mins	Biosorption	Binds with Hydroxyl, amino, methyl, carboxyl groups	(Guarín-Romero et al., 2019)
	Entero morpha intestinalis	28 ± 2°C, 9.92 pH, and 200 rpm.	93.38%	38.5 mins	Biosorption	-	(Hamouda et al., 2020)
	Cladophora glomerata	25 °C, 2 pH, and 200 rpm.	66%	1 h	Biosorptiom	-	(Al-Homaidan et al., 2018)
	Pseudochlorella pringsheimii and Chlorella vulgaris	27 ± 1 °C, 35 $\mu mol m^{-2} s^{-1}$.	80%	1 day	Bioaccumulation	Lipid accumulation increases	(Saranya & Shanthakumar, 2019)
	Oedogonium westi	18±1°C, 4 pH, natural light (12:12h).	93%	7 days	Biosorption	-	(Shamshad et al., 2016)
	Cystoseira	Cr(III): 4.5 pH	*C.barbata*: Cr(IV): 90.38% ;	Cr(III): 30-180 mins.	Biosorption	Alginic acid carboxyl groups	(Yalçın & Özyürek,

	barbata and Cystoseira crinita	and Cr(VI): 2 pH	Cr (III): 81.6% *C.crinite:* Cr(III): 82.86% ; Cr (IV): 89%	Cr(IV): 2-24 h.		for the retention of meta.	2018)
Cobalt	*Chlorococcum humicola*	25 ± 2 °C, 2000–3000 lx (14L: 10 D).	44%	-	Biosorption	Interact with hydroxyl, thiol or carboxyl surface active functional groups	(Borah et al., 2020)
Copper	*Desmodesmus sp. AARLG 074*	25 ◦C, 4 & 6 pH, 75 µmol photon/m2/s (White light).	80%	168 h	Biosorption & Bioaccmulation	Impact glutathione, galactose, arginine and proline metabolites.	(Buayam et al., 2019)
	Chlorophycea e spp.	23 ± 2ºC, 7 pH, 12L:12D (1000 µE/m^2/s) and 250 rpm.	88%	10 mins	Bioaccumulation	Effects Metabolites.	(Saavedra et al., 2018)
	Ulva lactuca	20 ± 2 °C, 7.8 pH, 12L:12D (natural Light).	86%	12 h	Biosorption and bioaccumulation	No effect of salinity.	(Henriques et al., 2019)
Iron	*Chlorococcum humicola*	25 ± 2 °C, 2000–3000 lux (14L: 10D).	74.47%	6 days	Biosorption	-	(Borah et al., 2020)
Lead	*Gelidium amansii*	45°C and 4.5 pH.	100%	2-24 h.	Biosorption	Binding sites: like hydroxyl, carboxyl, amino and sulfhydryl	(El-Naggar et al., 2018)
	Oedogonium westi	18±1°C, 5 pH, Natural light (12:12h).	61-96%	7 days	Biosorption	-	(Shamshad et al., 2016)
	Phormidium ambiguum	20 ± 1°C, 30 mE/m^2 /s, (16L:8D), 130 M/min.	70%	-	Biosorption and Bioaccumulation	Inhibitory effects towards growth.	(Shanab et al., 2012)
	Porphyra leucosticta	30°C, 8 pH, and 150 rpm.	Synthetic wastewater: 90% real	2 h	-	Bioremediation increases with increase in contact time.	(Ye et al., 2015)

Journal Pre-proof

			industrial waste water: 95%				
Manganese	Chlorella vulgaris	23 ± 2ºC, 7 pH, 12L:12D (1000 µE/m^2/s) and 250 rpm.	99.4%	3h	-	-	(Saavedra et al., 2018)
	Ulva lactuca	20 ± 2 °C, 7.8 pH, 12L:12D (natural Light).	74%	12 h	Biosorption and bioaccumulation	No effect of salinity.	(Henriques et al., 2019)
Mercury	*Ulva lactuca*	20 ± 2 °C, 7.8 pH, 12L:12D (natural Light).	98%	12 h	Biosorption and bioaccumulation	No effect of salinity.	(Henriques et al., 2019)
	Phormidium ambiguum	20 ± 1°C, 30 mE/m^2 /s, (16L:8D), 130 M/min	97%	-	Biosorption and Bioaccumulation	Inhibitory effects towards growth.	(Shanab et al., 2012)
Nickel	*Durvillaea antarctica*	20 ± 1°C, 5 pH and 160 rpm.	Adsorption capacity: 32.85 mg/g	240 mins.	Biosorption	Binds with Hydroxyl, amino, methyl, carboxyl groups	(Guarín-Romero et al., 2019)
	Oedogonium westi	18±1°C, 5 pH, Natural light (12:12h).	59-89% nickel	7 days			(Shamshad et al., 2016)
ZINC	Chlorophyceae *spp.*	23 ± 2ºC, 5.5 pH, 12L:12D (1000 µE/m^2/s) and 250 rpm.	91.9%	3h	-	-	(Saavedra et al., 2018)
	Fucus vesiculosus	20± 2ºC and 150 rpm.	-	-	Biosorption	Binding sites: Alginate and cellulose (Carboxyl and hydroxyl group)	(Brinza et al., 2020)

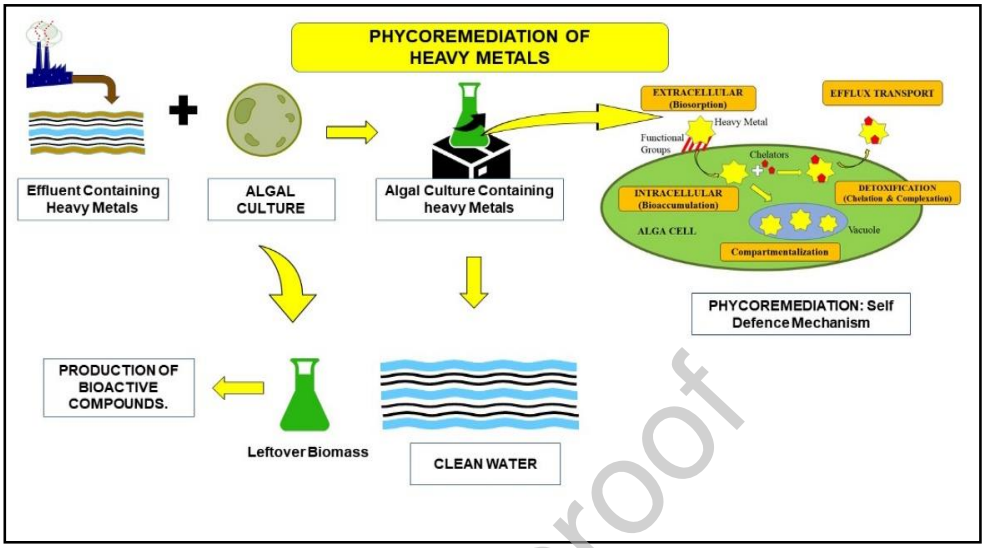
Journal Pre-proof

Table 2 Different value-added products along with the heavy metals abated by respective algae

Algal strain	By product	Productivity	Heavy metals abated	Reference
Chlamydomonas reinhardtii	Biofuel (Biodiesel, Biohydrogen) and therapeutic proteins (UV-protectant metallothionein, antibody and anti-cancer proteins)	-	Ar	(Scranton et al., 2017a)(Ramírez-Rodríguez et al., 2019)
Chlorella vulgaris	Biodiesel	Lipid: 0.2g/g	Mn, Ar, B, Cu and Zn	(Saavedra et al., 2018) (Mohammady et al., 2015)
Chlorococcum humicola	Bioethanol	520g/kg	Co, Cr and Fe	(Borah et al., 2020) (Patidar et al., 2017)
Desmodesmus sp.	*Biodiesel*	-	Cu	(Buayam et al., 2019) (Nzayisenga et al., 2020)
Durvillaea antarctica	Alginate and as biofertilizer.	-	Ni and Cr	(Guarín-Romero et al., 2019) (Mansilla et al., 2012)
Gelidium amansii	Bioethanol	-	Pb	(El-Naggar et al., 2018)(Park et al., 2012)
Phormidium sp.	Gold nanoparticles	-	Hg, Pb and Cd	(Shanab et al., 2012)(Chaudhary et al., 2020)
Porphyra sp.	Vitamin E, vitamin C, phytic acid and used in cosmetic industry.	-	Pb and Cd	(Salehi et al., 2019) (Ye et al., 2015)
Scenedesmus almeriensis	Lutein	8.54 mg/g	Ar, B, Cu, Mn and Zn	(Molino et al., 2019) (Saavedra et al., 2018)
Ulva lactuca	Fermentable Sugar and Silver Nanoparticles	Sugar yield: 155 mg/g dry biomass	Ar, Mn & Hg	(Henriques et al., 2019)(Turner et al., 2012) (Z. Z. Wu et al., 2018)

Journal Pre-proof

Graphical Abstract



Journal Pre-proof

Education and Information Technologies
https://doi.org/10.1007/s10639-022-11276-1

An alumni-based collaborative model to strengthen academia and industry partnership: The current challenges and strengths

Ruchika Malhotra[1] · **Massoud Massoudi**[1] · **Rajni Jindal**[2]

Received: 10 January 2022 / Accepted: 4 August 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Industry-academia collaboration (IAC) has long been a source of discussion in various ways. A vital first step toward closing the gap is fostering alumni engagement in industry and academia. Industry-academia collaborations are like earth and water, and they live in harmony. They cannot exist apart from one another. The collaborative effort between industry and academia promotes improvement and innovation in technology and innovative research and enables the university to be more relevant to the industry. This article proposes a collaborative approach based on alumni influences and uses secondary data on the influences of alumni in engineering to enrich the collaboration between industry and academia. The industry and academia utilize the alumni's potential to facilitate communication and offer their students and staff facilities. The research findings can be used by policymakers to innovate in the field of engineering and to strengthen the partnership between industry and academia in society. The model enables academia and industry to utilize the alumni strengths in enriching the collaboration; it also enhances academic research output, recruitment, fundraising, marketing, product promotion, and internship opportunities for students via alumni active participation. We propose some significant agendas for future work.

Keywords Industry-academia collaboration · Alumni association · Internship and research · Alumni engagement · Academia-alumni

✉ Ruchika Malhotra
ruchikamalhotra2004@yahoo.com

Massoud Massoudi
massoud.massoudi@hotmail.com

Rajni Jindal
rajnijindal@dce.ac.in

[1] Software Engineering Department, Delhi Technological University, Delhi, India

[2] Computer Science & Engineering, Delhi Technological University, Delhi, India

Published online: 15 August 2022

1 Introduction

A university establishes physical and virtual contact with students when they enroll as freshmen and continues throughout their time on campus. A university has been involved with students from the time they enroll until they graduate from college. When learners graduate and begin their careers, an alumni network is the only organization that can stay in touch with them. An alumni network will serve as a communication channel between the university and its graduates (Vittavas et al., 2015). Alumni are former students, faculty, and staff of a university, college, or school. Academia advancement describes the links between the institution and community organizations, such as alumni, who contribute to the university's vision and mission (Varsity, 2012).

The collaborative effort between industry and academia promotes improvement and innovation in technology and innovative research and enables the university to be more relevant to the industry. Alumni can serve as great role models for students and are often accepted. Alumni who return to help the university provide legitimacy and a reason for being a member of a successful institution, having breathed the same air and encountered similar obstacles as the students. Alumni experience in time management, financial planning, self-discipline, character development, and leadership development might be more readily received as counsel and inspiration by students. Alumni may aid in building confidence, increasing motivation, and instilling the appropriate culture consistent with the university's intended message to its students in this manner (Unisel, 2015).

Collaborative work has become a rapidly growing focal point as a necessary tool for overcoming various industrial difficulties (Demircioglu & Audretsch, 2019). In recent years, universities and industries have been increasingly aware of the significant benefits of aligning and integrating academic affairs, student affairs, and alumni services around shared aims and objectives. Indeed, there are various possibilities for student affairs professionals to collaborate with alumni affairs counterparts. Both sections aim to improve the institution's image and the experience of people who interact with it. They seek to accomplish these goals by collaborating with basically the same community but at various stages of their relationship with the university. When this type of harmonic collaboration is established and sustained, the advantages to student and alumni affairs are substantial (Singer & Hughey, 2002).

University education is the highest degree of academic instruction. It provides students with extensive knowledge and research skills and the opportunity to engage with and establish relationships with their classmates, instructors, and other academic people. These relationships are critical for graduates since they may inspire new initiatives and collaborations after their degrees. As a result, industries and academics promote these relationships (Sauermann & Stephan, 2010; Gama & Garcia, 2021). Furthermore, academia and industry have already highlighted the necessity and benefits of the partnership. Many industries continue to struggle with encouraging and establishing partnerships (de Wit-de Vries et al., 2019; Awasthy et al., 2020). The questions like 1. How does interacting

with alumni meet the university's needs? 2. How does engagement with the university meet the requirements of alums? have to be cleared. If not, it will be purposeless and a waste of energy and resources. There are increased attempts to foster cooperation, and the current research identifies numerous different types of collaboration. Nevertheless, industries continue to face challenges in identifying best practices for establishing and maintaining partnerships. These best practices have increased the necessity for establishing specific norms and principles that will govern the beginning, execution, and effectiveness of a partnership (Awasthy et al., 2020).

Collaboration between academic and industrial sectors is becoming more widely recognized to enhance innovation through information exchange. The concept is stated by a large increase of research that looks at the subject from many angles. On the other hand, this information collection is splintered and lacks a detailed picture. Another significant gap between industry and academics is the incompatibility of education systems with industry requirements; educational institutions cannot be consigned to the background. They form the foundation for all of society's activities (Ankrah & AL-Tabbaa, 2015). There is a considerable discrepancy between what students are taught and what industry expects in today's education system, resulting in decreasing production (Rishika, 2021).

The present industry challenge is that the gap between industry expectations and institutional outcomes continues to widen daily (Natthawat et al., 2018). Alumni input should be the key resource for determining current industry requirements. They contribute to the growth of the institute after receiving their degree. Alumni can help bridge the gap between academia and industry by organizing conferences, expert talks, guest faculties, and student workshops on cutting-edge research (Kakad & Dhage, 2021). Alumni passion and dedication may be vital for fundraising, engagement, marketing, and product promotion. To reap those rewards, academia must keep in contact with their learners even after they graduate. This article focuses on the alumni's important role in industry-academia collaboration. However, the term can also refer to students, instructors, and other people connected to a school or program. Alumni associations are official, organized networks of people who have a common programmatic or educational relationship (Campbell & Baxter, 2019).

Alumni affairs at the majority of colleges and universities employ a variety of tactics aimed at assisting them in becoming constructive allies with other campus groups and industries in their shared objective of enhancing and enriching the college experience. Hence, we propose a collaborative model for enhancing industry-academia collaboration through alumni interactions. We discuss a variety of programs that successfully connect academia and industry through alumni for mutual benefit. Finally, we explore the educational benefits that result from education institutions sponsoring and supporting these types of effective collaboration. The remainder of the paper is structured as follows: In Section II, the collaboration between academia and industry is addressed; in Section III, alumni responsibilities in industry-academia collaboration are discussed. Section IV introduces the proposed model, sections V discusses the study validation, and sections VI and VII conclude the study with a discussion and conclusion.

2 Academia and industry collaboration

Academia is a generic term that refers to the community of learners, educators, and researchers involved in higher education and research. The term originates in Greek and refers to a larger body of knowledge, development, and transfer across generations. Academia in universities unquestionably has the research capability, motivation, and expertise necessary to impart the core of education to broad groups of a community via training and programs (Nangia & Pramanik, 2011). The industry encompasses all economic activities that result in the production of commodities or services. It is a link in a chain - from raw materials to finished product, then to the service sector, and finally to research & innovation. Industries, the countries they operate, and their economies are all entangled in a complicated web of interconnectedness (Sen et al., 2015). Academic institutions are under immense pressure due to rising stakeholder and learner expectations (Sliep & Marnewick, 2020). Industry and academia require close relationships to enrich the current activities by supporting each other. In a recent paper (Weatherby et al., 2022), the need for stronger collaboration between academics who design educational technologies, educators and students (who utilize educational technology), and academics who know how to do research on its efficacy and effectiveness was discussed. The five levels of closeness between academia and industry are outlined in Fig. 1. (source:(Wohlin, 2013b)). Undoubtedly, commerce is one of the most novel partnerships in human history (i.e., A partnership in which one party purchases and another offers a service or product). One of the key advantages of this type of collaboration is that what each partner gets and provides is visible. Partnerships between industry and academia based on commerce and in which the research party plays the role of a typical customer may be beneficial in certain circumstances for productive research initiatives. The clarity of both parties' give-and-take promises is a significant advantage over many other types of cooperation. Each industry–research collaboration raises ethical difficulties,

Fig. 1 Five levels of IA closeness

exacerbated when the industry member is uninformed of the process's principal aim. The primary issue may be how much loss to participants is incurred, including time-wasting (Jørgensen, 2017). Industry and academia collaborate for various reasons: industries profit from highly trained people resources such as researchers and students; they benefit from access to technology and information, and they may utilize sophisticated research infrastructure (Ankrah & AL-Tabbaa, 2015). According to some studies, university research contributes up to 10% of digital products or technologies (Bekkers & Freitas, 2008). Academics need industry collaboration for more financing, access to industry technology, and patenting revenue. Indeed, collaboration with industry has become an integral aspect of academia funding and contributions to Research & development activities in the higher education sector from international organizations and technology firms and constitute a significant source in many nations (Barnes et al., 2002). In addition, collaborations with other sectors give the opportunity of resolving complex problems in academics in terms of resources, the transmission of knowledge, and the development of individuals' skills (Bidandi et al., 2022).

Because of these consequences and their economic significance, it is critical to ensure the successful management of industry and academia collaborations to maximize the benefits for both parties (Rybnicek & Königsgruber, 2019). Approximately twenty years ago, design-related disciplines facilitated the introduction and rise of the user experience field, which is evident in both academia and industry through the increasing number of user experience journals, conferences, as well as the expanding job market.

In comparison to its efforts to conceive and investigate user experience, academia has fallen behind in offering user experience education programmes (Süner-Pla-Cerdà et al., 2021). Additionally, the model proposed in this study will significantly strengthens this collaboration. It depends on how much academia and industry value collaboration and partnerships to improve academic and industrial services to people and society. This is possible if both organizations use alumni as a medium of connection and establish strong ties with alumni. This interaction provides academia and industry many opportunities for professional and interpersonal improvements.

In the no-contact level of closeness, academia and industry work independently and solve their current challenges. At this level, academia is busy with their research work and publishing for academic improvements (Fig. 2).

As a result, when there is no interaction between industry and academia, the below activities do not improve, and the gap widens (Jørgensen, 2017).

- There is a lack of incentives.
- There is no foundation for a center of excellence and significance.
- There is no diffusion of knowledge via peer-to-peer interaction.
- Alumni are not involved as mentors to students.
- No recruitment of top-notch faculty
- There is no provision for establishing interface structures.
- Wastage of academia-industry research and development lab
- There is no cell dedicated to entrepreneurship development and technology incubation.

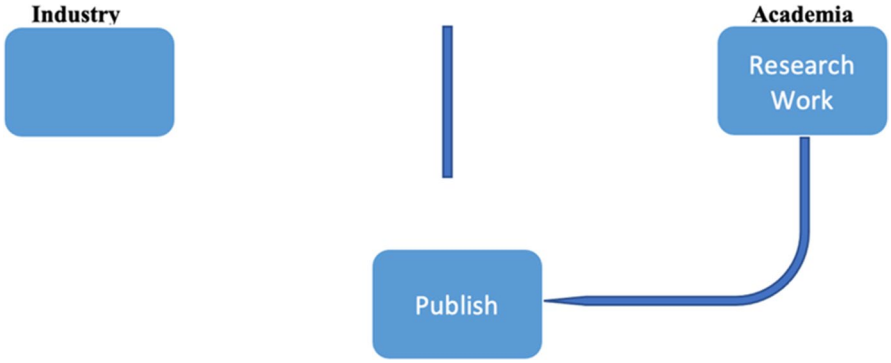


Fig. 2 Industry and academia with no contact

- There is no mechanism to facilitate technology transfer from the laboratory to market access.
- There is no chance of improvement in the transfer of technology toward commercialization.
- There is no establishment of a research park to promote enterprise growth.
- There will be no venture capital funds established to support innovative entrepreneurship.

Rumor, or hearsay, is the second level of proximity between industry and academia. Here, academia is made aware of industry-wide challenges. They hear about problems and defects in the systems available in the industry. Academia is trying to develop new methods and approaches to solve the current issues available in the systems in industries. Academia addresses the industry's concerns, as illustrated in Fig. 3, develops solutions, and publishes the results in journals and conferences; however, academia does not share the outcomes with the industry. Hence, sharing the solutions with industries, academia requires cost for the solutions developed and

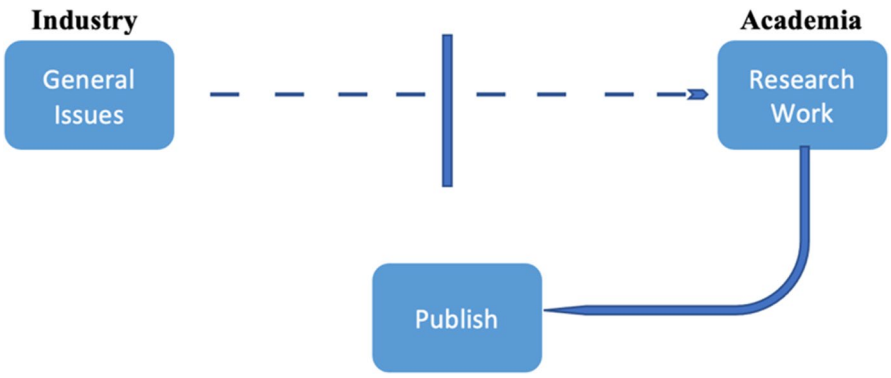


Fig. 3 Hearsay between the industry and academia

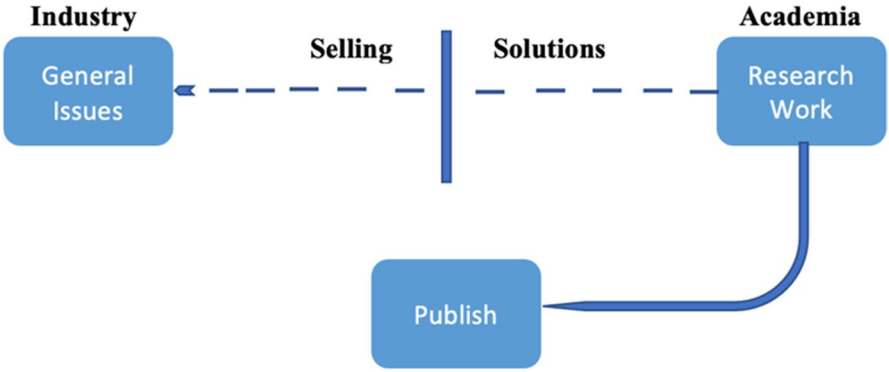


Fig. 4 Sales pitch between industry and academia

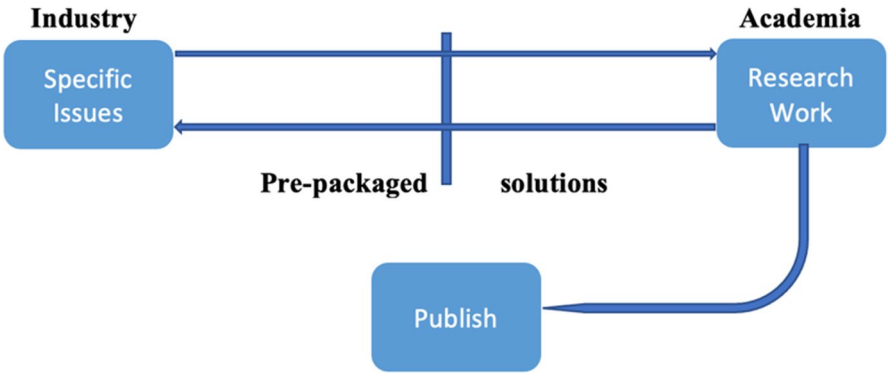


Fig. 5 Offline mode interaction between IAC

researched by its researchers. To do so, the third level of closeness comes into the picture.

According to (Wohlin, 2013b), the commercial pitch is the third level of proximity between academia and industry. Academia visits the industry to discuss their challenges; they knock on their doors and make sales pitches to sell solutions and approaches. Since academia will not offer the answers and procedures to the industry for free, they will need to sell their approach to the industry. This level attempts to push and pull academia and industry together by identifying industrial difficulties and developing solutions to them. Nevertheless, it's all about cost vs. quality; the more you pay, the better solutions you receive. As per Fig. 4, academia develops solutions to industry-wide problems and monetizes the findings.

The fourth level of closeness is the "offline stage." There is contact between industry and academia at this level, as shown in Fig. 5. The industry has specific challenges that they share with academia, and academia collaborates with the industry offline, conducting research on the challenge and developing a list of pre-packaged solutions and submitting it to the industry.

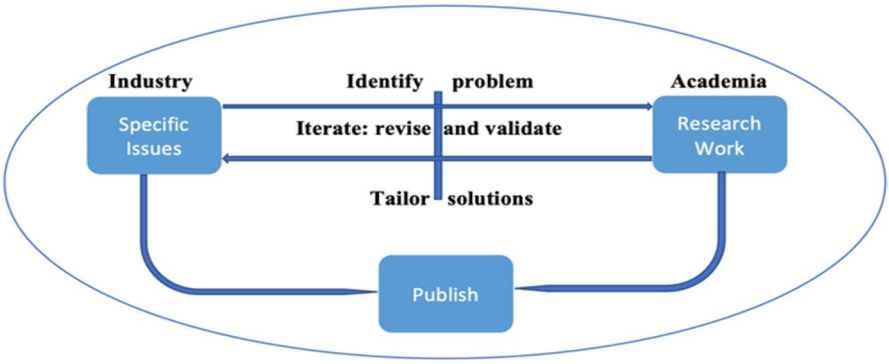


Fig. 6 Teamwork between IAC

Fig. 7 Potential students' advantages

The final level of proximity is the united team or so-called one team, as shown in Fig. 6. At this level, industry and academia collaborate closely. Several industry-specific issues will be discovered together by academia and industry. Firstly, they tailor solutions, revise and validate them, and publish the outcome. Finally, they act to be one team with robust ideas and potential.

2.1 Current industry challenges

Researchers and practitioners published numerous studies on the vital benefits of industry-academia collaboration during the last decade. Such cooperation is widely acknowledged to benefit both technology firms' interests. Additionally, many engineering students and educators profit from partnerships between academia and industry since they exchange information and acquire experience (Raghavan & Towhidnejad, 2006). Students' potential advantages of the collaboration between academia and industry are listed in Fig. 7 (Harrison, 1997; Kornecki et al., 1997, 2003). Academic interaction and collaboration with alumni are still unusual for a variety of reasons, including the following: a) alumni do not believe universities have provided them with valuable information and incentives, b) alumni are tracked solely to obtain funding contributions to universities, c) academic institutions generally treat productive and experienced alumni members better than current and much less productive alumni delegates, and d) the relationship between academia and alumni remains strained. Additionally, there are barriers to academia's engagement with their alumnus (Natthawat et al., 2018).

Undoubtedly, educational and industrial systems form the backbone of the technological world. Numerous ontologies have been established in education (Chin & Chang, 2011). Every day, the gap between industry expectations and institute outcomes widens. Alumni input should be a primary source of information for determining current industry requirements. They contribute to the growth of the institute after completing their degree. The distance between industry and academia can be bridged through hosting conferences, expert talks, and student workshops on current developments using alumni power (Kakad & Dhage, 2021). Garousi (Garousi et al., 2016a) conducted a systematic study of industry-academia collaboration and identified various current barriers and best practices. Table 1 summarizes multiple obstacles that industries face and their associated best practices.

2.2 Current industry strengths

The industry is perpetually engaged in targeted development in the term's broadest sense. The industry's preferred time frames for engagement with academia are immediate, and investment is motivated by efforts that deliver quantifiable outcomes (Sen et al., 2015). Five possible industry strengths for engineering education are presented in Fig. 5, and they might be considered industry strengths for academia (Letizia & Morasca, 2006).

To begin with, teaching courses by industrial experts enables software engineering instructors to understand better what matters in the industry and how it may be taught in college lectures. There are distinctions between the instructor's interaction with industry professionals who attend classes and the instructor's relationship with university students. Industry experts have various requirements, which are frequently rather precise, whereas university students have more generalized objectives (Fig. 8).

In addition, alumni, particularly teacher education candidates, serve as an ideal bridge between an educator and industry. Numerous software specialists have completed their first industrial-grade project while working on their dissertations, which may have assisted them in establishing a career in an industrial company. After graduation, both alumni and instructors' perspectives will undoubtedly shift. The alumni may realize that the instructor's career responsibilities extend way beyond the dissertation topic or the undergraduate courses delivered by the teacher. These responsibilities enable communication to be more expansive and adapt its scope over time as technology, and industrial demands evolve.

On the other hand, instructors' research and teaching activities are less constrained in scope and time than practitioners' industry activities. As a result, alumni may consult their advisers and instructors for longer-term viewpoints and alternatives to their current techniques. From the teacher's perspective, a researcher's close contact with alumni may facilitate and encourage communication, providing the instructor and researcher with a somewhat privileged channel to the software industry through which they can distinguish trending topics and patterns for use in their academic and professional programs (Lethbridge, 1998; Letizia & Morasca, 2006).

Table 1 Industries' various challenges and best practices

References	Challenges	Best practices	References
(Runeson, 2012; Garousi et al., 2016a)	Inadequate research relevance • The research findings are entirely irrelevant for practice. • Researchers do not comprehend the pertinent issues from the standpoint of industry.	• Require long-term plans • Allow researchers quick and regular access (makes solutions pragmatic and feasible). • Organize workshops and seminars (provides access to industry-related issues).	(Osterweil et al., 2008; Martínez-Fernández & Marques, 2014; Krishnan et al., 2009; Garousi et al., 2016b)
(Morries et al., 1998)	Inadequate training, expertise, and competencies • a lack of education in software engineering. • Lack of information on the researcher's business context and technologies utilized in practice.	• Industry becomes more receptive to involvement in education. • Changes in the content of courses in academia. • Utilize the case study technique (case studies are effective tools for disseminating knowledge and expertise). • The researcher and the industry site should be co-located.	(Krishnan et al., 2009; Svenér, 2014; Gorschek et al., 2006; Rico et al., 2021)
(Kaindl et al., 2002)	Lack of enthusiasm or dedication • Failure to provide access and time when promised.	• Need for top teams, as well as their mindsets. • Maintain frequent contact via meetings (provides internet connectivity.	(Grünbacher & Rabiser, 2013) (Runeson, 2012)
(Petersen & Engström, 2014; Sandberg et al., 2011a)	Incompatibility of industry and academia • Distinct terminologies and modes of communication. • Diverse modes of communication and orientations of the flow of information.	• Prior solid experience requires interaction. • Interact directly with practitioners while gathering data.	(Grünbacher & Rabiser, 2013; Runeson & Minör, 2014)

Table 1 (continued)

References	Challenges	Best practices	References
(Martínez-Fernández & Marques, 2014; Morries et al., 1998; Garousi et al., 2019)	Factors relating to people and organizations • Inflexibility and resistance to change. • Instability and discontinuity within the organization. • Working as team • Issues of technical and organizational nature	• Utilize the case study method (just for experimental purposes). • Ascertain management involvement on the industry side (minimizes the organization's influence). • Awareness • Communication	(Baldassarre et al., 2013; Runeson, 2012; Rico et al., 2021; Rybnicek & Königsgruber, 2019)
(Morries et al., 1998; Wohlin, 2013a; Krishnan et al., 2009)	Concerns about management • Complexity in establishing clear and realistic project aspirations and goals. • Commitment to make a significant time/effort investment.	• Appropriate presentation and communication skills on the part of researchers in early sessions. • Demonstrate the benefits of the research solution to the industrial partner (time-saving solutions are valuable).	(Sandberg et al., 2011b; Enoiu & Čaušević, 2014)

Fig. 8 The five main strengths of an industry that academics benefits from

Another pertinent point is that industrial experts may bring knowledge and experience to regular college classes. Experienced software engineers often have a different viewpoint on real-world obstacles and challenges than those provided in academic settings. Professionals have a distinct perspective from lecturers, enabling students to grasp the significance of previously missed themes. For instance, many software engineering students believe the most crucial step is the coding phase. This phase creates the final product, which is also used in the early software classes and often entails software development. It may be difficult for educators to convince students of the value of other stages (e.g., early stages). Thus, educators may find it challenging to convey the industrial aspects of software development in general. Learners frequently overlook administrative and organizational concepts. Hence, students may discover professionals with a real-world experience more engaging than academics. They can justify judgments made on non-technical grounds. The learners are often highly responsive and eager to learn from a representation of the real world. Software professionals may educate in several ways, from making brief presentations to teaching whole courses. Software professionals need to comprehend the relationship between their lectures and other disciplines covered in the curriculum. It's self-evident that software developers should contact instructors to determine which topics can be taught successfully, to what extent, and results. Additionally, software professionals may use this expertise to develop and teach courses inside their organizations (Letizia & Morasca, 2006).

Consequently, software engineering has frequently followed an ideological path, unlike many other engineering fields. Industrial companies would undoubtedly gain a competitive advantage if they could get innovative ideas and approaches from academic theoretical research and data on their effectiveness or appropriate usage. In exchange, software industries will receive an assessment of methods and techniques that may be applied. The laboratory experiment results can serve as a pilot study for further investigations in the industrial organization. Additionally, the line between students and software professionals is increasingly blurred in numerous instances. Students work as consultants or in part-time or full-time occupations. Professionals return to college to further improve their knowledge (Carver et al., 2003). Finally, the literature on software engineering education has several articles about team project courses and the disciplines required to work on software teams. We believe it is incredibly advantageous for students to learn software engineering ideas and methods through traditional lectures, academic exercises, and hands-on experience that closely resembles industrial practice. Thus, industry relevance is needed for software engineering projects, even if presenting a full-fledged industrial project to students is rarely feasible.

On the other hand, industrial firms may provide substantial subsets of application demands that may be used during college sessions, such as homework assignments, class projects, and demonstrations. Jaccheri in (Jaccheri, 2001) discusses an example of this form of industrial contact. Another example is the Norwegian University of Science and Technology (Andersen et al., 1994). A company, therefore, becomes the client, requiring some amount of participation, which varies from circumstance to scenario.

3 Alumni roles in industry-academia collaboration

Although the term can also refer to students, instructors, and other people connected to a school or program, alumni associations are official, organized networks of people who have a typical programmatic or educational relationship (Campbell & Baxter, 2019). A supportive, engaged alumni network is critical to an institution's success. After a graduate leaves a university, their perception of the institution becomes static. Rather than that, they should be kept informed of the university's progress frequently. Both the university and the alumni profit from alumni connections. Here, higher education alumni relations professionals examine the significance of an engaged alumni network and how the market is changing. Researchers and practitioners must work collaboratively throughout the partnership process (Wohlin, 2013b). Bachelor's degrees typically take three to four years to complete. If students study without creating industry contacts, their qualifications are questioned. Few college students get work experience. Graduates pursue new careers. Those without industry contacts may struggle to find work (AlGhamdi, 2019).

Several alumni associations collaborate with university career services and give information on incoming students and alumni job placement opportunities. Certain universities offer alumni and students access to online job recruitment and search tools (such as MonsterTRAK.com). These technologies enable job seekers to search a database of national and international openings (Singer & Hughey, 2002). A characteristic frequently connected with career success is identifying a mentor who will aid students in making the transition from college to the workplace. Alumni are uniquely qualified for this position. They can draw on their expertise, assisting students in exploring the world of employment, working, and finding jobs. Indeed, there are many advantages to alumni working directly with students, mainly as they make their initial career choices. Alumni are uniquely positioned to provide practical instruction and support to students pursuing degrees in fields related to alumni careers. Along with the personal satisfaction that comes from contributing to another person's development, participating alumni get to witness students who may one day be candidates for positions within their own business. Additionally, numerous alumni organizations provide work shadowing and mentorship programs, summer job placement options, and internship opportunities for currently enrolled students and assist alumni employers in connecting with prospective employees (Letizia & Morasca, 2006).

Consequently, alumni are likely to be the university's most ardent supporters. "*An engaged alumni network enables the university to benefit from graduates'*

strengths and knowledge by aiding our students, the institution, and one another," said Molly Southwood, the University of Bath's Head of Alumni Relations. Alumni are our most loyal members and finest advocates, but only if we keep them informed and engaged. A brief role of the alumni network is outlined in Fig. 9. Alumni with remarkable talents and knowledge may share their expertise with current students via speeches, newsletters, and even social media. In some cases, alumni may even help students get jobs and start their careers.

Alumni can serve as international ambassadors for the college in various situations. Precisely, at the City University of London. "*More than 45% of City's students are international*," stated Celia Enyioko-Hanniford, the director of the alumni communications office. They assist us in demonstrating to prospective students what life is like in London and persuading them to study with us. In today's competitive job market, these upgrades give learners an edge.

Regarding the program's outcomes, we have maintained our status as one of the top universities in the UK for graduate employability and beginning wage. Alumni are often generous with fundraising efforts due to their loyalty and gratitude to the university. "*Financial contributions enable us to give life-changing scholarships and bursaries to worthy students who might otherwise be forced to stop their studies. They also allow the city to provide students with contemporary teaching and research facilities*". While an engaged alumni network serves the university, it also benefits the alumni. It's been satisfying, said Kenneth Okoroafor (BSc Economics and Accountancy 2005) in an interview for our annual alumni magazine. "I've learned a lot about myself and my career. As graduates matured and new alumni emerged, the market naturally altered. What worked before may no longer work for a new generation of graduates". Approaches may need revision and modification (Cannon, 2015). Alumni Affairs is actively involved in several mutually beneficial university initiatives on behalf of students. These collaborative initiatives include programs intended to improve the overall quality of student life, strategies for orienting and welcoming new students to campus, and activities geared at attracting and maintaining students (Singer & Hughey, 2002). Several academic institutions have established initiatives in collaboration with admissions departments that directly involve alumni recruiting prospective students. In (Hanson, 2000), Hanson discovered that many variables correctly predict which alumni are most inclined to advocate for their institutions. These criteria include the organization's status, social identification, years since graduation, and respect for alumni leaders. Most schools and universities aggressively recruit potential students through institutional rituals

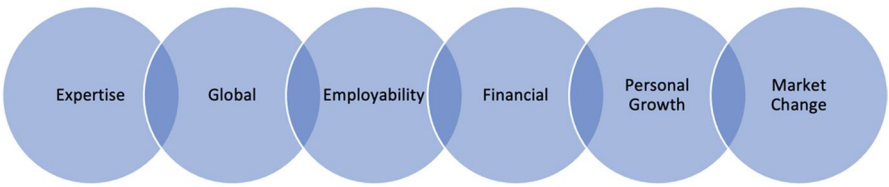


Fig. 9 Role of alumni network in IAC

and traditions. Parents, staff, and alumni frequently accept and actively promote these institutional views. The key to any alumni recruitment plan is to identify elements that contribute a solid tie to the institution, its present students, and its alumni (Kellogg, 1998).

4 Proposed model

The proposed model is a collaborative model designed based on the alumni influence to enrich the industry and academia collaboration. The term "alumni engagement" encompasses a broad range of activities (Di Battista et al., 2014; Farr-Wharton et al., 2018; Snijders et al., 2019). Fostering interaction is a worldwide concern and critical for industrialized countries' educational settings (Coates, 2010). Numerous universities and alumni organizations organize random events to stimulate the alumni's interest. Following the success of an alumni engagement activity, the activity is repeated a few times until alumni discover the activity is no longer exciting, and involvement levels quickly decrease to a bare minimum. Alumni groups and universities are rarely clear on how various events may fit into a broader engagement model and how various ideas can be created to fulfill alumni needs. Most universities are unaware of the relevance of alumni engagement and, as a result, do not invest time or money in it. There is apathy in academia about this vital work. A model for alumni engagement would aid in comprehending the different components of connection and lend precision to this involvement (Parthasarathy, 2018). We developed an innovative engagement model depicted in Fig. 10 to strengthen industry-academia collaboration. This model outlines alumni roles and responsibilities toward

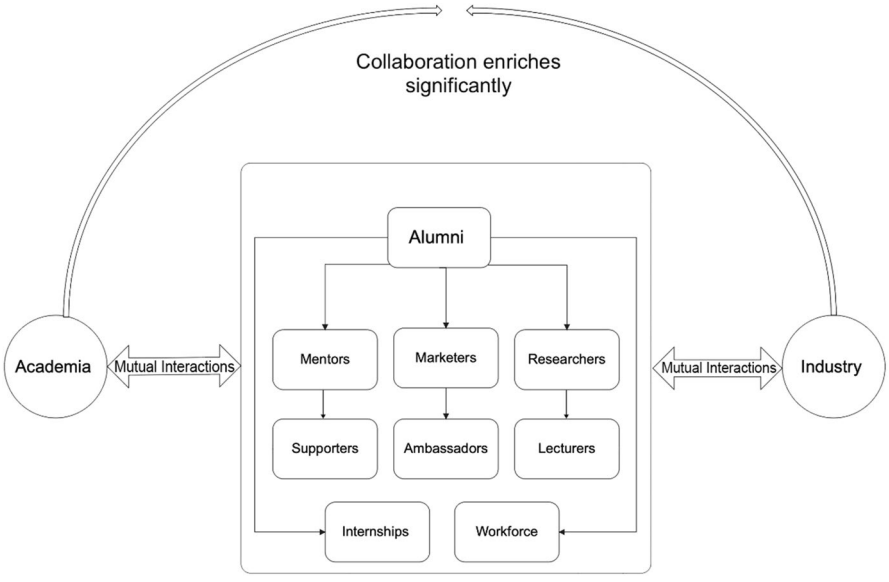


Fig. 10 Alumni engagement model to strengthen industry-academia collaboration

industry and academia. According to the model, alumni can contribute to academia and industry, including mentors, marketers, researchers, supporters, ambassadors, lecturers, internship providers, and a productive workforce.

The model aims to strengthen and establish relationships between academia and industry through mutual interaction with alumni. Alumni are past students of academia and industries' current and future workforce if these two organizations maintain positive relationships with alumni and pave the way for alumni to engage in academia for mentoring and providing internships for current and future students of academia. This positive relationship results in the establishment of a favorable channel of interaction. Additionally, the concept stipulates that industry and academia must maintain a mutual relationship with alumni. Close interaction with academia encourages alumni to of more supportive of their academia. The list of potential alumni as per the above Fig. 10 is detailed as follows:

a. Alumni as mentors: Mentoring research with undergraduate university students has. However, most of the attention has been on instructor-to-student mentoring (Lunsford, 2011; Underhill, 2006). Other forms of mentorship with university students, such as mentoring interactions with alumni, have received limited attention, resulting in a knowledge gap on the issue. Conner in (Conner, 2015) did qualitative research in his Ph.D. dissertation to better understand the expectations and experiences of undergraduate learners mentored by alumni through a mentorship program organized by the university career center. According to the study's findings, most participants sought help from their alumni mentors regarding their job and academic endeavors. Interactions between university students and their mentors occasionally included university-specific information that the students considered valuable. According to interviewees, students' interactions with their alumni mentors were constructive since they gained industrial and educational knowledge. Students reported enhanced confidence, increased career clarity, and decreased stress in the present due to their increased understanding of potential future career opportunities as a result of their mentoring interactions with alumni. There are several factors why students seek mentorship from alumni. Some were investigating majors and wanted to know what they might do with them, while others had particular career paths and desired help in that subject. As a student interested in software engineering, he especially sought a mentor with that expertise.
b. Alumni as marketers: Academic institutions rely on alumni associations for ongoing support and should consequently cultivate, retain, and develop alumni partnerships (McAlexander & Koenig, 2001). Accordingly, the collaboration is typically accomplished through the alumni association, which serves as a marketing tool and spokesperson for the university and its alumni. Certain colleges (Rust, 2015) are using an alumni organization as a marketing tool to foster positive relationships with their alumni. A good relationship is primarily dependent on recognition. Humans feel a sense of belonging when they are in a relationship where both sides recognize the value of respecting one another's duties. The institution will then work to leverage the support of existing students, alumni, and the broader community for its initiatives. The alumni network is an efficient tool for publi-

cizing, organizing, and managing specific initiatives (Born, 2014). Universities solicit contributions from alumni continuously. That demands a mindset in which the university welcomes the client, values the client, and understands that the better they treat their customers, the higher their service quality. Universities adopt mass marketing techniques at considerable expense to maintain or expand their market share (Shaik, 2005). Marketing budgets continue to grow, with institutions extensively spending on an integrated marketing mix that targets potential learners. However, determine the potential benefit of using relationship management principles to graduate marketing.

Shaik and McAlexander (McAlexander & Koenig, 2001; Shaik, 2005) concur that alumni serve as a marketing instrument and spokesperson for the university's public and new and prospective students to encourage recruitment and enrollment. At the same time, learners can participate in alumni events, purchase university-branded products, participate in registrations when deciding to continue their education, and contribute to the university's income.

c. Alumni as researchers: It has been asserted that the strength of alumni-alma mater strong connections has an important influence on alumni interaction. A long-lasting relationship motivates alumni to engage in more helpful behavior toward their alma mater (Heckman & Guskey, 1998). This type of assistance may involve alumni being generous with contributions, advising individuals orally to participate in their alma mater, joining the advisory board, and offering managerial support (Leslie & Ramey, 1988; Weerts & Ronca, 2008). Is there a close connection between alumni relationships and faculty behavior in collaborative research? This question is answered by (Li et al., 2015) by assessing the collaboration proclivities of alumni lecturers and non-alumni instructors. Finally, Alumni can be a valuable asset to academia when conducting applied research (Boardman & Ponomariov, 2007).

d. Alumni as supporters: Alumni are considered academia's most passionate supporters. Hence, an active and responsible alumni platform enables the institution to use alumni' abilities and knowledge to benefit its students, institutions, and others. A robust alumni network can be highly beneficial to academia. Alumni can aid their alma mater in various ways, including providing non-monetary or financial support. As a result, they form an intriguing and significant entity (Iskhakova et al., 2017). Student engagement, in particular, refers to the learners' feelings of attachment to the academic institution in which they are admitted, as exhibited by their attitudes or behavioral personality traits (Hennig-Thurau et al., 2001). Two interconnected factors serve as the foundation for that attachment: attitude and behavior (Snijders et al., 2019). Attitudes toward alumni attachment can be stated as "a desire to provide monetary assistance, a desire to stay connected with the academic, an interest in receiving academic news, and a commitment to join the alumni organization" (Iskhakova et al., 2016). Behavioral alumni attachment is frequently manifested by financial assistance (Snijders et al., 2019). Most academia maintains these resources to grow their programs, facilities, and scholarship opportunities. Thus, without raising funds, a university will be unable to maintain and expand consistently with the needs of new learners. Hence, funding departments seek to solicit funds from alumni. Alumni can be

beneficial supporters of academia via mentoring, sponsoring, providing internships for students, hosting events, conducting conferences, and working with students' research activities in academia.

e. Alumni as ambassadors: Universities that attract more international students certainly produce a greater number of global ambassadors. Similarly, both internal and overseas graduates serve as ambassadors for the institution and the country they studied. They acquire knowledge, curricula, systems, and policies and partially or entirely implement the experiences and strategies gained in that country when they return to their home country (Lomer, 2017). Lomer also stated in his research finding: That overseas students are willing to take part in both the culture and politics of the United Kingdom and professional training. They intend to participate in the academic and local communities, developing an understanding of "British values" and knowledge and awareness of life in the United Kingdom. They are supposed to have a positive experience, form friendships and "enduring bonds" with faculty and students, and acquire political sympathy for the UK's interests. Students are expected to return home once their studies are completed to exert influence on their homecoming, or at the very least once they reach the pinnacle of their careers. They must maintain contact with the United Kingdom and their host institution. These expectations are replicated and spread through policy, alumni, and organizational networks.
f. Alumni as lecturers: Numerous universities ask alumni to lecture, engage in panel discussions, and share their experiences with students as guest faculty members. Such initiatives benefit students and alumni by providing essential learning opportunities and reconnecting them with their alma mater (Singer & Hughey, 2002). Students need to learn about occupations and the business world during their academics and understand the required skills to thrive in the labor market to direct their studies meaningfully. We believe that allowing students to engage with alumni from their university who work in their field of study would drive them to stay involved in their studies and will aid them in making career decisions (Kailanto et al., 2012). Instructors from the engineering field have emphasized the need to incorporate "real world" experiences into the classroom. Specific individuals have lobbied for strategies to facilitate student contact with alumni and other professionals, like mentoring programs (Sharp, 2004). Alumni as guest faculty are one of the productive alternatives for academia to recruit them to teach software engineering students industry-based experiences.
g. Alumni as internship providers: Alumni are the most loyal and enduring members of academia; they have a strong connection to their alma mater. When they work in industry, they can play a critical role in bridging the gap between academia and industry. They have direct access to all recruiting and internship opportunities in their industry. Additionally, when an institution's relationship with its alumni is positive, it gives the university more internship possibilities for current students through the assistance of alumni who work at the company. A characteristic frequently connected with career success is identifying a mentor who will assist learners in making the transition from academia to the workplace. Alumni are uniquely qualified for this position. They can depend on their expertise to help

college students explore the employment market, network, and find jobs, among other things.

h. Nevertheless, alumni have several advantages working directly with students, mainly when they make their initial career choices. Alumni are in a role that can provide feasible education and support to students pursuing degrees in fields related to alumni careers. Along with the personal satisfaction that comes from supporting another person's development, participating alumni get to witness students who may one day be candidates for positions within their institutions. Numerous alumni organizations support work monitoring and mentoring programs, summer job placement opportunities, and internship opportunities for recently enrolled students and assist alumni employers in connecting with potential job seekers (Singer & Hughey, 2002).
i. Alumni as workforce: This alumni feature encompasses most of the primary factors that alumni demonstrate in their roles as marketers, mentors, lecturers, ambassadors, and supporters of their university and industry. Apart from demographic challenges, the mindsets of today's workforce affect firms and their ability to stay effective. Alumni are the only workforce available to academia and industry, and they are the driving factor behind the increased collaboration between the two. With regard to academia, alumni as a workforce can significantly contribute to the advancement of academia in local and international communities by assisting students in acquiring the necessary knowledge to meet industry needs, conducting applied research, and attending industry-sponsored events.

As a result, the preceding model is comprehensive and is implemented to ascertain alumni's academic and industrial needs. Industry and academia need to determine the alumnus's requirements, roles, and possibilities. What do they consider necessary? And how do they contribute to the enrichment of academia-industry collaboration? Additionally, they must strengthen their engagement with alumni and guarantee that alumni continue to engage with academia frequently.

5 Study validation

Effective communication is essential in any collaborative effort. Industries that develop software and services continuously attempt to enhance their software engineering capabilities to compete effectively. This entails gaining access to individuals with appropriate expertise and expanding the company's present depth of knowledge. Academics strive to provide both parts of competence by producing software developers who are employable in the industry and doing scientific studies that contribute to industry knowledge and enhance industrial processes (Rico et al., 2021). Alumni are academia's most devoted and abiding members; they have a sense of belonging with their institution. They can play a significant role in reducing the gap between academics and industry when they work in the industry. They get immediate access to their industry's employment and internship possibilities.

Additionally, when a university maintains a strong relationship with its alumni, it provides additional internship opportunities for current students through the support

of alumni who work at the organization. A trait usually associated with job success is locating a mentor who can aid learners in transitioning from academia to the workplace. Alumni are unmatched in their qualifications for this role. They can rely on their skills to assist college students in navigating the job market, networking, and finding work, among other things (Singer & Hughey, 2002). The majority of colleges are unaware of the importance of alumni engagement, and as a result, they do not allocate time or money in this area. In academics, this important endeavor is met with disinterest. A model for alumni interaction would enable us to understand the various components of connection and enhance the precision of this engagement (Parthasarathy, 2018).

Subsequently, we presented a collaborative approach based on alumni influence to improve industry-academia collaboration in engineering for the first time. The model's uniqueness is that it leverages alumni potential in industry and academia to strengthen partnerships and provide multiple services to academia. Including technical research assistance to students, guest faculty lectures, sharing practical knowledge required in the industry, facilitating internship opportunities for students in academia, conducting workshops on research and innovation, and many more academic opportunities.

6 Discussion

We devised an innovative model centered on the alumni role to bolster industry-academia collaboration. This model is critical for connecting academia with alumni and improving academia's connection with industry. Alumni with extraordinary abilities and experience may share their knowledge and expertise with current students through speeches, newsletters, and even social media. Alumni may even assist students in obtaining jobs and launching their careers in some instances. The model in Fig. 10 illustrates the most typical duties and responsibilities of alumni, which may strengthen the connection between industry and academia if the alumnus maintains a strong association with academia throughout and after their academic career. In maintaining the industry and academia collaboration, alumni play a vital role in keeping this engagement strong and long-lasting. A supportive, engaged alumni network is essential to the success of any academia. Following graduation from a university, a graduate's perception of the institution becomes fixed. Rather than that, they should be kept regularly informed of the university's progress. Alumni connections benefit both the university and the alumni. Here, alumni relations experts in higher education discuss the value of an engaged alumni network and how the market is changing (Wohlin et al., 2012). On the one hand, some academic institutions do not pay enough attention to their alumni and do not give opportunities for them to share their skills and knowledge with students. On the other hand, numerous colleges invite alumni to speak as guest faculty members, participate in panel discussions, and share their experiences with students. These programs serve students and alumni by offering necessary educational opportunities and reuniting them with their alma mater (Singer & Hughey, 2002). The innovative model, which is based on alumni involvement in industry and academia, assists all academic institutions in

developing a solid connection with their alumni and strengthening their collaboration with industry, ensuring that their students graduate with appropriate industry experience. In connection with the industry-academia collaboration, there is a need to have close relationships to enrich the current activities by supporting each other. Wohlin in (Wohlin, 2013b) developed a five-tiered system of proximity between industry and academia and insisted on frequent contact between the two. The more in touch the two organizations are, the more collaborative they will become. Industry and academia can achieve more fruitful results if they collaborate as a cohesive team. In this case, industry and academia collaborate to develop solutions for industry-specific challenges, which can be done in academia. Additionally, he stated that the close relationship between industry and academia results in academia providing a sales pitch to find a solution and approach to the issues encountered in industry and requesting payment in exchange for the solution they provided to specific issues encountered in industrial systems. The industry and academia suffer from challenges and issues. According to (Natthawat et al., 2018), the current industry challenge is the gap between industry expectations and institution outcomes continues to increase daily. Alumni feedback should be the primary source of information used to determine current industry requirements. After earning their degree, they contribute to the institute's growth. The gap between industry and academics can be overcome by using alumni influence to hold conferences, expert speeches, and student workshops on current advances (Kakad & Dhage, 2021). To overcome current industry difficulties, several best practices are presented to the current industry challenges (Garousi et al., 2016a) for overcoming current industry difficulties. Firstly, there is a need to set long-term goals and provide researchers with timely and regular access to develop pragmatic and feasible solutions to the industry's challenges. Exposes researchers seminars provides researchers with exposure to industry issues and challenges. Another essential practice is to keep in touch frequently via meetings. To conclude, researchers' adequate presentation and communication skills and ability to demonstrate the benefits of research solutions to industrial partners are significant. Academic engagement represents an important way academic knowledge is transferred into the industrial domain; many companies consider it significantly more valuable than licensing university patents (Cohen et al., 2002). The five primary strengths of an industry that benefit academia are summarized in Fig. 8. Firstly, the industry acts as a learner; it sends software specialists to academia to offer courses to students to gain a better grasp of industry priorities. Additionally, the industry serves as alumni—the ideal connection between educator and industry. From a teacher's perspective, a researcher's close relationship with alumni may encourage and enable communication, providing the educator and researcher with a somewhat privileged channel to the software industry to identify trending topics and patterns for inclusion in their academic and professional lives programs (Letizia & Morasca, 2006). Finally, the innovative model, built on alumni influence, requires a supporting and involved alumni network to ensure the success of institutions and industries. They can use their knowledge to assist learners in learning about the world of work, working, and finding employment. Furthermore, alumni engaging directly with students have a distinct advantage, mainly as students make their initial career choices. Alumni are uniquely positioned to teach and mentor students seeking

degrees in subjects linked to alumni careers. Several studies (Assefa, 2016; King & Persily, 2020; Sen et al., 2015) have discussed the relationship between industry and academia using various best practices and theoretical models. However, neither study has proposed a model based on alumni influence to strengthen industry-academia collaboration. Our study proposes an innovative model based on alumni influences to strengthen the linkages between industry and academia. The approach attempts to strengthen and establish links between academia and industry through mutual involvement with alumni. If these two organizations maintain strong relationships with alumni and facilitate alumni engagement in academics through mentoring and internship opportunities for current and future students of academia. This positive relationship results in forming a channel of communication that is beneficial. Additionally, the approach establishes a mutual interaction between industry and academics and their alumni. Alumni who have a close relationship with their universities are more likely to support them.

7 Conclusion

To conclude, it is essential to consider the mutual benefits of a partnership between industry and academia. Both need a high level of commitment and collaboration. Their relationships are similar to those of earth and water. They are unable to work independently. However, in most cases, the collaboration between industry and academia witnesses a considerable gap. What academia teaches does not meet industry requirements, and academia does not produce what industry requires. This rift demands a bridge to interact and collaborate between these two organizations effectively. The five levels of closeness in section two are fruitful factors that bridge the gap between industry-academia and the organization required to enrich their ties as one team and work collaboratively.

The model's findings indicate that when alumni participate in various activities such as delivering workshops, giving lectures to students, mentoring students' research projects, and providing internship chances, the partnership between industry and academia is strengthened. The academia-alumni connection promotes alumni to be more supportive of their academic institutions. Similarly, Academic connection and collaboration with alumni remain uncommon for a variety of reasons, among them are the following: a) alumni believe universities have failed to provide them with valuable information and incentives, b) alumni are tracked solely to obtain funding contributions to universities, c) academic institutions generally treat productive and experienced alumni members better than current and much less productive alumni delegates, and d) the relationship between academia and alumni continues to be strained. Additionally, we summarized some of the best practices available in industry and academics for overcoming existing barriers in Table 1. The collaborative model proposed in this study significantly enriches the current status of industry-academia collaboration through the use of alumni influence. Alumni play a critical role in fostering collaboration by participating actively in academic activities. For example, guest faculty teaching, mentoring students, conducting seminars, working on applied research, marketing, and assisting academia in various ways.

The primary model focuses on mutual interaction between industry and academia. That is possible if both organizations use alumni as a medium of connection and establish strong ties with alumni. This interaction provides academia and industry many opportunities for professional and interpersonal improvements. The academia focusing and insisting on linking their alumni with current students enriches their reputation as quality educational institutions. In a nutshell, a supportive and engaged alumni network is significant for the success of the industry and academia. Future researchers may expand on this concept by conducting a complete survey of numerous organizations and academic institutions for an impactful outcome and further strengthen the relationship between industry and academia.

Data availability As no dataset were generated or analysed during the course of this study, data sharing is not applicable to this article.

Declarations

Conflict of interest None.

References

AlGhamdi, R. (2019). Fostering information technology students' internship program. *Education and Information Technologies, 24*(5), 2727–2739. https://doi.org/10.1007/s10639-019-09891-6

Andersen, R., Conradi, R., Krogstie, J., Sindre, G., & Sølvberg, A. (1994). Project courses at the NTH: 20 years of experience. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 750 LNCS* (Issue June 2014). https://doi.org/10.1007/bfb0017613

Ankrah, S., & AL-Tabbaa, O. (2015). Universities-industry collaboration: A systematic review. *Scandinavian Journal of Management, 31*(3), 387–408. https://doi.org/10.1016/j.scaman.2015.02.003

Assefa, A. A. (2016). University- industry linkage practices , determinants and challenges theoretical and empirical article review: Lessons for effective and successful collaboration. *International Journal of Research in Management, Economics and Commerce*, *6*(03), 1–16. www.indusedu.org

Awasthy, R., Flint, S., Sankarnarayana, R., & Jones, R. L. (2020). A framework to improve university–industry collaboration. *Journal of Industry-University Collaboration, 2*(1), 49–62. https://doi.org/10.1108/jiuc-09-2019-0016

Baldassarre, M. T., Caivano, D., & Visaggio, G. (2013). Empirical studies for innovation dissemination: Ten years of experience. *ACM International Conference Proceeding Series*, 144–152. https://doi.org/10.1145/2460999.2461020

Barnes, T., Pashby, I., & Gibbons, A. (2002). Effective university - industry interaction: A multi-case evaluation of collaborative R&D projects. *European Management Journal, 20*(3), 272–285. https://doi.org/10.1016/S0263-2373(02)00044-0

Bekkers, R., & Freitas, B. M. I. (2008). *Analysing knowledge transfer channels between universities and industry: To what degree do sectors also matter?* Research Policy. https://econpapers.repec.org/RePEc:eee:respol:v:37:y:2008:i:10:p:1837-1853. Accessed 12 Oct 2021

Bidandi, F., Anthony, A. N., & Mukong, C. (2022). Collaboration and partnerships between south African higher education institutions and stakeholders: Case study of a post-apartheid university. *Discover Education*, *1*(1). https://doi.org/10.1007/s44217-022-00001-2

Boardman, P. C., & Ponomariov, B. L. (2007). Reward systems and NSF University research centers: The impact of tenure on university scientists' valuation of applied and commercially relevant research. *The Journal of Higher Education, 78*(1), 51–70. https://doi.org/10.1080/00221546.2007.11778963

Born, S. (2014). *University Alumni Relations Involve Networking and Maintaining Links, Alumni Portal,* From http://www.alumniportal-deutschland.org/en/study-continuing-education/study-training/article/university-alumni-relations-networking-maintain-ing-links.html. Accessed 2 Oct 2020

Campbell, A. C., & Baxter, A. R. (2019). Exploring the attributes and practices of alumni associations that advance social change. *International Journal of Educational Development, 66*(October), 164–172. https://doi.org/10.1016/j.ijedudev.2018.10.003

Cannon, T. (2015). *The importance of the alumni network*. Red Brick. https://www.redbrickresearch.com/2015/10/29/the-importance-of-the-alumni-network/. Accessed 5 Nov 2020

Carver, J., Jaccheri, L., Morasca, S., & Shull, F. (2003). Issues in using students in empirical studies in software engineering education. *Proceedings - International Software Metrics Symposium*, *2003-Janua*, 239–249. https://doi.org/10.1109/METRIC.2003.1232471

Chin, K. L., & Chang, E. (2011). A sustainable ICT education ontology. *IEEE International Conference on Digital Ecosystems and Technologies, 5*(June), 350–354. https://doi.org/10.1109/DEST.2011.5936612

Coates, H. (2010). Development of the Australasian survey of student engagement (AUSSE). *Higher Education, 60*(1), 1–17. https://doi.org/10.1007/s10734-009-9281-2

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and impacts: The influence of public research on industrial R&D. *Management Science, 48*(1), 1–23. https://doi.org/10.1287/mnsc.48.1.1.14273

Conner, K. (2015). Expectations and experiences of undergraduate students who participated in an alumni mentoring program. In *ProQuest Dissertations and Theses*. https://ezproxy.lib.uwm.edu/login?url=https://search.proquest.com/docview/1705827201?accountid=15078%0Ahttp://uwi-primoalma-prod.hosted.exlibrisgroup.com/openurl/MIL/mil_services_page?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&genre. Accessed 10 Nov 2021

de Wit-de Vries, E., Dolfsma, W. A., van der Windt, H. J., & Gerkema, M. P. (2019). Knowledge transfer in university–industry research partnerships: A review. *Journal of Technology Transfer, 44*(4), 1236–1255. https://doi.org/10.1007/s10961-018-9660-x

Demircioglu, M. A., & Audretsch, D. B. (2019). Public sector innovation: The effect of universities. *Journal of Technology Transfer, 44*(2), 596–614. https://doi.org/10.1007/s10961-017-9636-2

Di Battista, S., Pivetti, M., & Berti, C. (2014). Engagement in the university context: Exploring the role of a sense of justice and social identification. *Social Psychology of Education, 17*(3), 471–490. https://doi.org/10.1007/s11218-014-9255-9

Enoiu, E. P., & Čaušević, A. (2014). Enablers and impediments for collaborative research in software testing: An empirical exploration. *WISE 2014 - Proceedings of the 2014 ACM International Workshop on Long-Term Industrial Collaboration on Software Engineering, Co-Located with ASE 2014*, 49–54. https://doi.org/10.1145/2647648.2647655

Farr-Wharton, B., Charles, M. B., Keast, R., Woolcott, G., & Chamberlain, D. (2018). Why lecturers still matter: The impact of lecturer-student exchange on student engagement and intention to leave university prematurely. *Higher Education, 75*(1), 167–185. https://doi.org/10.1007/s10734-017-0190-5

Gama, D., & Garcia, R. (2021). The role of academic relations of former graduate students in university-firm collaboration: Evidence from Brazil. *The Journal of Technology Transfer.* https://doi.org/10.1007/s10961-021-09881-2

Garousi, V., Petersen, K., & Ozkan, B. (2016a). Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review. *Information and Software Technology, 79*(July), 106–127. https://doi.org/10.1016/j.infsof.2016.07.006

Garousi, V., Petersen, K., & Ozkan, B. (2016b). Challenges and best practices in industry-academia collaborations in software engineering: A systematic literature review. *Information and Software Technology.* https://doi.org/10.1016/j.infsof.2016.07.006

Garousi, V., Pfahl, D., Fernandes, J. M., Felderer, M., Mäntylä, M. V., Shepherd, D., Arcuri, A., Coşkunçay, A., & Tekinerdogan, B. (2019). Characterizing industry-academia collaborations in software engineering: Evidence from 101 projects. *Empirical Software Engineering.* https://doi.org/10.1007/s10664-019-09711-y

Gorschek, T., Garre, P., Larsson, S., & Wohlin, C. (2006). A model for technology transfer in practice. *IEEE Software, 23*(6), 88–95. https://doi.org/10.1109/MS.2006.147

Grünbacher, P., & Rabiser, R. (2013). Success factors for empirical studies in industry-academia collaboration: A reflection. *2013 1st International Workshop on Conducting Empirical Studies in Industry, CESI 2013 - Proceedings*, 27–32. https://doi.org/10.1109/CESI.2013.6618466

Hanson, S. K. (2000). *Alumni Characteristic that Predict Promoting and Donating to Alma Mater: Implications for Alumni Relations*. 119.

Harrison, J. V. (1997). Enhancing software development project courses via industry participation. *Software Engineering Education Conference, Proceedings*, 192–203. https://doi.org/10.1109/sedc.1997.592454

Heckman, R., & Guskey, A. (1998). The relationship between alumni and university: Toward a theory of discretionary collaborative behavior. *Journal of Marketing Theory and Practice, 6*(2), 97–112. https://doi.org/10.1080/10696679.1998.11501799

Hennig-Thurau, T., Langer, M. F., & Hansen, U. (2001). Modeling and managing student loyalty: An approach based on the concept of relationship quality. *Journal of Service Research, 3*(4), 331–344. https://doi.org/10.1177/109467050134006

Iskhakova, L., Hilbert, A., & Hoffmann, S. (2016). An integrative model of alumni loyalty—An empirical validation among graduates from German and Russian universities. *Journal of Nonprofit and Public Sector Marketing, 28*(2), 129–163. https://doi.org/10.1080/10495142.2015.1006490

Iskhakova, L., Hoffmann, S., & Hilbert, A. (2017). Alumni loyalty: Systematic literature review. *Journal of Nonprofit and Public Sector Marketing, 29*(3), 274–316. https://doi.org/10.1080/10495142.2017.1326352

Jaccheri, M. L. (2001). Software quality and software process improvement course based on interaction with the local software industry. *Computer Applications in Engineering Education, 9*(4), 265–272. https://doi.org/10.1002/cae.10000

Jørgensen, M. (2017). Working with Industry: Stories of Successful and Failed Research-Industry Collaborations on Empirical Software Engineering. *Proceedings - 2017 IEEE/ACM 5th International Workshop on Conducting Empirical Studies in Industry, CESI 2017*, 46–52. https://doi.org/10.1109/CESI.2017.3

Kailanto, M., Silius, K., Tervakari, A. M., Marttila, J., Tebest, T., & Huhtamaki, J. (2012). Alumni stories and students' commitment to studies. *IEEE Global Engineering Education Conference, EDUCON*. https://doi.org/10.1109/EDUCON.2012.6201166

Kaindl, H., Brinkkemper, S., Bubenko, J. A., Farbey, B., Greenspan, S. J., Heitmeyer, C. L., do Prado Leite, J. C. S., Mead, N. R., Mylopoulos, J., & Siddiqi, J. (2002). Requirements engineering and technology transfer: Obstacles, incentives and improvement agenda. *Requirements Engineering, 7*(3), 113–123. https://doi.org/10.1007/s007660200008

Kakad, S., & Dhage, S. (2021). *Building Alumni Ontology to bridge Industry-Institute Gap using Protégé 5.5. 5*, 1–6. https://doi.org/10.1109/incet51464.2021.9456325

Kellogg, K. O. (1998). An analysis of the collaborative programming between student affairs and alumni relations professionals at select post-secondary institutions in Missour. In *Dissertation: Vol. Ph.D. Thes* (Issue Structural Biology and Molecular Biophysics, University of Pennsylvania, PA, USA.).

King, G., & Persily, N. (2020). A new model for industry-academic partnerships. *PS - Political Science and Politics, 53*(4), 703–709. https://doi.org/10.1017/S1049096519001021

Kornecki, A. J., Hirmanpour, I., Towhidnajad, M., Boyd, R., Ghiorzi, T., & Margolis, L. (1997). Strengthening software engineering education through academic industry collaboration. *Software Engineering Education Conference, Proceedings*, 204–211. https://doi.org/10.1109/sedc.1997.592455

Kornecki, A. J., Khajenoori, S., Gluch, D., & Kameli, N. (2003). On a partnership between software industry and academia. *Software Engineering Education Conference, Proceedings*, *2003-Janua* (January), 60–69. https://doi.org/10.1109/CSEE.2003.1191351

Krishnan, P., Ross, K., & Pari-Salas, P. (2009). Industry academia collaboration: An experience report at a small University. *Proceedings - 22nd Conference on Software Engineering Education and Training, CSEET 2009*, 117–121. https://doi.org/10.1109/CSEET.2009.22

Leslie, L. L., & Ramey, G. (1988). Donor behavior and voluntary support for higher education institutions. *The Journal of Higher Education, 59*(2), 115–132. https://doi.org/10.1080/00221546.1988.11778318

Lethbridge, T. C. (1998). The relevance of software education: A survey and some recommendations. *Annals of Software Engineering, 6*(1–4), 91–110. https://doi.org/10.1023/a:1018917700997

Letizia, J., & Morasca, S. (2006). *On the importance of dialogue with industry about software engineering education*. 5. https://doi.org/10.1145/1137842.1137845

Li, F., Miao, Y., & Yang, C. (2015). How do alumni faculty behave in research collaboration? An analysis of Chang Jiang scholars in China. *Research Policy, 44*(2), 438–450. https://doi.org/10.1016/j.respol.2014.09.002

Lomer, S. (2017). Influence: A political rationale and international alumni as ambassadors. *Palgrave Studies in Global Higher Education*, 99–125. https://doi.org/10.1007/978-3-319-51073-6_5

Lunsford, L. G. (2011). Psychology of mentoring: The case of talented college students. *Journal of Advanced Academics, 22*(3), 474–498. https://doi.org/10.1177/1932202X1102200305

Martínez-Fernández, S., & Marques, H. M. (2014). Practical experiences in designing and conducting empirical studies in industry-academia collaboration. *2nd International Workshop on Conducting Empirical Studies in Industry, CESI 2014 - Proceedings*, 15–20. https://doi.org/10.1145/2593690.2593696

McAlexander, J. H., & Koenig, H. F. (2001). University experiences, the student-college relationship, and alumni support. *Journal of Marketing for Higher Education, 10*(3), 21–44. https://doi.org/10.1300/J050v10n03_02

Morries, P., Masera, M., & Wilikens, M. (1998). Requirements engineering and industrial uptake. *Proceedings of IEEE International Symposium on Requirements Engineering*, 130–137. https://doi.org/10.1109/ICRE.1998.667818

Nangia, V. K., & Pramanik, C. (2011). Towards an integrated model for academia-industry interface in India. *World Academy of Science, Engineering and Technology, 73*(1), 333–342.

Natthawat, R., Sinthupinyo, S., & Chandrachai, A. (2018). An innovation model of alumni relationship management: Alumni segmentation analysis. *Kasetsart Journal of Social Sciences, 39*(1), 150–160. https://doi.org/10.1016/j.kjss.2017.02.002

Osterweil, L. J., Ghezzi, C., Kramer, J., & Wolf, A. L. (2008). Determining the impact of software engineering research on practice. *Computer, 41*(3), 39–49. https://doi.org/10.1109/MC.2008.85

Parthasarathy, S. (2018). *The Alumni Engagement Model*. LinkedIn. https://www.linkedin.com/pulse/alumni-engagement-model-parthasarathy-s. Accessed 27 Sept 2021

Petersen, K., & Engström, E. (2014). Finding relevant research solutions for practical problems-the serp taxonomy architecture. *WISE 2014 - Proceedings of the 2014 ACM International Workshop on Long-Term Industrial Collaboration on Software Engineering, Co-Located with ASE 2014*, 13–20. https://doi.org/10.1145/2647648.2647650

Raghavan, J., & Towhidnejad, M. (2006). *2006–424 : CHALLENGES / ISSUES IN A INDUSTRY-ACADEMIC Challenges in an Industry-Academic Collaboration*.

Rico, S., Bjarnason, E., Engström, E., Höst, M., & Runeson, P. (2021). A case study of industry–academia communication in a joint software engineering research project. *Journal of Software: Evolution and Process*, *July*, 1–19. https://doi.org/10.1002/smr.2372

Rishika, R. (2021). *A Study on Challenges to Bridge the Gap between Academia and Industry – With Reference To Post Graduation Courses from Industrial and Students Perspective*. *March*.

Runeson, P. (2012). It takes two to tango - An experience report on industry - Academia collaboration. *Proceedings - IEEE 5th International Conference on Software Testing, Verification and Validation, ICST 2012*, 872–877. https://doi.org/10.1109/ICST.2012.190

Runeson, P., & Minör, S. (2014). The 4+1 View Model of Industry–Academia Collaboration. *Proceedings of the 2014 International Workshop on Long-Term Industrial Collaboration on Software Engineering*, 21–24. https://doi.org/10.1145/2647648.2647651

Rust, A. A. (. B.). (2015). The role of students and alumni in relationship marketing: An exploratory study at the Cape Peninsula University of Technology, South Africa. *International Journal of Educational Sciences, 9*(1), 37–46. https://doi.org/10.1080/09751122.2015.11890292

Rybnicek, R., & Königsgruber, R. (2019). What makes industry–university collaboration succeed? A systematic review of the literature. *Journal of Business Economics, 89*(2), 221–250. https://doi.org/10.1007/s11573-018-0916-6

Sandberg, A., Pareto, L., & Arts, T. (2011a). Agile collaborative research: Action principles for industry-academia collaboration. *IEEE Software*. https://doi.org/10.1109/MS.2011.49

Sandberg, A., Pareto, L., & Arts, T. (2011b). Agile collaborative research: Action principles for industry-academia collaboration. *IEEE Software, 28*(4), 74–83. https://doi.org/10.1109/MS.2011.49

Sauermann, H., & Stephan, P. (2010). Twins or Strangers? Differences and similarities between industrial and academic science. *National Bureau of Economic Research*. https://doi.org/10.3386/w16113

Sen, S., Ganguly, S., & Sen, J. (2015). Bridging the Industry-Academia Gap and Knowledge Management: Need of Hour. *3rd South Asian Geoscience Conference & Exhibition*, *January*.

Shaik, N. (2005). Marketing distance learning programs and courses: A relationship marketing strategy. *Online Journal of Distance Learning Administration, 8*(2), 1–7.

Sharp, J. E. (2004). Work in progress: Alumni mentoring of engineers in a technical communication course. *Proceedings - Frontiers in Education Conference, FIE, 2*, 11–12. https://doi.org/10.1109/fie.2004.1408608

Singer, T. S., & Hughey, A. W. (2002). The role of the alumni association in student life. *New Directions for Student Services, 2002*(100), 51–68. https://doi.org/10.1002/ss.70

Sliep, C., & Marnewick, C. (2020). The quest in delivering quality IT services: The case of a higher education institution. *Education and Information Technologies, 25*, 4818–4844. https://doi.org/10.1007/s10639-020-10198-0

Snijders, I., Wijnia, L., Rikers, R. M. J. P., & Loyens, S. M. M. (2019). Alumni loyalty drivers in higher education. *Social Psychology of Education, 22*(3), 607–627. https://doi.org/10.1007/s11218-019-09488-4

Süner-Pla-Cerdà, S., Günay, A., Töre Yargın, G., & Ural, H. (2021). Industrial design students' perceptions towards a career in user experience field in Turkey. *International Journal of Technology and Design Education, 32*(3), 1895–1923. https://doi.org/10.1007/s10798-021-09666-6

Svenér, J. (2014). *Get the Cogs in Synch – Time Horizon Aspects of Industry – Academia Collaboration*. 25–28.

Underhill, C. M. (2006). The effectiveness of mentoring programs in corporate settings: A meta-analytical review of the literature. *Journal of Vocational Behavior, 68*(2), 292–307. https://doi.org/10.1016/j.jvb.2005.05.003

Unisel, alumni. (2015). *Role of alumni in university development*. https://uniselalumni.wixsite.com/uniselalumniconnect/single-post/2015/10/15/. Accessed 13 Aug 2021

Varsity, F. (2012). Role of alumni in university development. In http://erepository.uonbi.ac.ke (Vol. 4, Issue 1). http://erepository.uonbi.ac.ke/bitstream/handle/11295/62456/Roleofalumniinuniversityadvancement.pdf?sequence=1. Accessed 22 Jul 2021

Vittavas, R., Sinthupinyo, S., & Chandrachai, E. A. (2015). An innovation system that can quickly responses to the needs of students and alumni. *Procedia - Social and Behavioral Sciences, 182*, 645–652. https://doi.org/10.1016/j.sbspro.2015.04.801

Weatherby, K., Clark-Wilson, A., Cukurova, M., & Luckin, R. (2022). The importance of boundary objects in industry-academia collaborations to support evidencing the efficacy of educational technology. *TechTrends*, *0123456789*. https://doi.org/10.1007/s11528-022-00705-0

Weerts, D. J., & Ronca, J. M. (2008). Characteristics of alumni donors who volunteer at their alma mater. *Research in Higher Education, 49*(3), 274–292. https://doi.org/10.1007/s11162-007-9077-0

Wohlin, C. (2013a). Empirical software engineering research with industry: Top 10 challenges. *Proceedings of the First International Workshop on Conducting Empirical Studies in Industry (CESI)*, 43–46.

Wohlin, C. (2013b). Software Engineering Research under the Lamppost. *Icsoft*, IS–11.

Wohlin, C., Aurum, A., Angelis, L., Phillips, L., Dittrich, Y., Gorschek, T., Grahn, H., Henningsson, K., Kågström, S., Low, G., Rovegard, P., Tomaszewski, P., Van Toorn, C., & Winter, J. (2012). The success factors powering industry-academia collaboration. *IEEE Software.* https://doi.org/10.1109/MS.2011.92

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Analysis of municipal solid waste management strategies in the city of Greater Noida

V. Varna Vishakar [a], B. Dhilipkumar [b,*], Ayush Jain [a], Zohaib Ahmed Khan [a]

[a] Department of Civil Engineering, KIET Group of Institutions, Delhi-NCR, Meerut Road(NH-58), Ghaziabad, India
[b] Department of Civil Engineering, K. Ramakrishnan College of Technology, Samayapuram, Trichy, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Waste management
Waste administration framework
Central Pollution Control Board (CPCB)
Greater Noida Industrial Development Authority (GNIDA)
Hazardous waste
Infectious waste
Municipal Solid Waste (MSW)

ABSTRACT

India is a developing nation with growing economy, demonstrating high asset utilization and waste age especially in urban area. Metropolitan strong waste is arranged informally in the greater part of the urban Indian urban communities, breaking ecological pathways. This examination was focused at building up a stock of strong garbage removal offices in the up and coming significant metropolitan city Greater Noida, Uttar Pradesh by get-together and dissecting applicable information on essential and auxiliary assortment and removal frameworks for example per capita age, per capita removal, assortment productivity, assortment and capacity/office, transportation offices, treatment and removals and the significant preventions in Solid Waste Management (SWM). Strong Waste Management was investigated concerning assortment proficiency and isolation of waste.
© 2023 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Newer Engineering Concepts and Technology.

1. Introduction

Considering the trouble in obtaining land for building up squander the executives and removal offices, it is basic that the current dumpsites are updated to get present and future squanders. As aimless dumping over the dumpsites has been a typical practice, and tainting of the encompassing territories isn't exceptional, those dumpsites should be restored and upgraded to recuperate space for future wastes [2]. The restoration measures ought to incorporate measures to contain the contaminant relocation and where conceivable covering the base of the fill and structure it to get future squanders for a time of 20 years of additional through imaginative structures and all around arranged tasks and maintenance [3]. Waste generation is increasing continuously depending upon the population growth, on an average up to 2021 census data 41% of population increase that much amount of waste gets constantly increasing per year [5]. Analysis done by the Central Pollution Control Board for determining the Waste management strategies in the nation revealed that on an average in the major metropolitan cities 48 million tons Municipal solid waste is getting generated constantly [1]. In India, squanders are typically high in biodegradable issue and low in paper, metal and glass. The extents of the constituents additionally shift occasionally and here and there relying upon way of life, food propensities, way of life and level of business and modern activity [4].

Noteworthy increment has been seen in MSW age in India over the most recent couple of decades due to quick populace development, financial turn of events and change in ways of life, food-propensities, and expectations for everyday comforts and expanding commercialization in urban focuses. MSW is gathered and dumped on the recognized removal or unidentified dumping locales, which are typically the low-lying zones on the edges of the city [11]. Open dumping of trash encouraged genuine medical issues, ground water defilement by leachate creation and emanation of climatic contaminations. Evaluations given by Tata Energy Research Institute (TERI) for open-decline consuming, showed most elevated TSP heap of yearly emanation among PM10, SO2 and NOx [9]. The wellsprings of MSW (Municipal Solid Waste) are family units, workplaces, shops, lodgings and different organizations. Calorific estimation of the Indian strong waste has been accounted for to be somewhere in the range of 600 and 800 kcal/kg and the thickness of the loss somewhere in the range of 330 and 560 kg/m^3.Greater Noida likewise experiences the issue of

* Corresponding author at: Assistant Professor, K. Ramakrishnan College of Technology, Trichy, India.
E-mail addresses: varnavishakar@gmail.com (V. Varna Vishakar), dhilipkaan@gmail.com (B. Dhilipkumar), ayushjain2904@gmail.com (A. Jain), zohaib.khan@kiet.edu (Z. Ahmed Khan).

https://doi.org/10.1016/j.matpr.2022.07.419
2214-7853/© 2023 Elsevier Ltd. All rights reserved.
Selection and peer-review under responsibility of the scientific committee of the International Conference on Newer Engineering Concepts and Technology.

Please cite this article as: V. Varna Vishakar, B. Dhilipkumar, A. Jain et al., Analysis of municipal solid waste management strategies in the city of Greater Noida, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2022.07.419

open dumping in numerous areas which are not being recognized by the specialists and makes irritation to the individuals just as the environment [12–16]. Despite the fact that treating the soil has been the pervasive organic preparing rehearsed in India, there have been issues because of transportation, helpless acknowledgment by ranchers (might be a result of value concerns), promoting, cost and so forth. In larger part of urban focuses, squander is being discarded in low-lying areas [10]. The removal locales are chosen based on their nearness to the assortment zones and new removal destinations are ordinarily distinguished just when the current ones are filled. Much of the time, the waste is just dumped at such destinations and, aside from in the significant urban areas, pieces of machinery are once in a while utilized for compaction at the removal site [6]. Indeed, even in these urban areas, they are utilized essentially for leveling of the stored squander. Appropriate gauging, filling and soil layering are not rehearsed in numerous regions. Arrangements for leachate and gas control don't exist at numerous places [7,8]. A dirt spread is once in a while given, aside from at the hour of definite conclusion of the site. The greater part of the removal locales are unfenced and the waste picking is normally stylish, presenting issues in the activity of the destinations. It is a typical practice to light a fire on the dumpsite by the cloth pickers either to lessen the threatening flies and volume or smell and encourage squander picking. The motivation behind this examination was to comprehend and break down the civil strong waste administration. The investigation was attempted to contemplate the quality and attributes of strong waste, comprehend the spatial example of waste administration and distinguish fitting system by following the issue identified with civil strong waste administration with certain proposals for better administration.

2. Methodology

Keeping in view, a severe need of dependable information book on assortment and age of MSW amounts, transportation and treatment and removals offices in Greater Noida, a vital arrangement for breaking down the status of strong garbage removal offices is created. Squander age being straightforwardly corresponding to populace size; the urban areas were grouped concerning the populace size (Greater Noida population evaluation 2011). A point by point investigation of civil profile of Uttar Pradesh was made and a study poll was planned assessing the means of strong waste administration for example age, assortment, treatment and removal. Field trips and postal studies were incorporated. In view of the civil profile, the chose review locales were grouped by their populace size. The underneath Fig. 1 shows the Municipal Solid Waste composition in an overall scenario perspective.

More prominent Noida is situated in Gautam Budh Nagar region of Uttar Pradesh state India. More prominent Greater Noida is around 30 km southeast of New Delhi, 30 km northwest of the region central command, Greater Noida location availability around 447 km from the northwest towards Lucknow which is the state capital.

Greater Noida lies on the banks of Yamuna River which is located northwest to the Major metropolitan city New Delhi. On the other side urban communities of Uttar Pradesh, Noida, and Ghaziabad are located in the north east direction where in the south east direction towards of Hindon River.

More prominent Greater Noida situated under the territory of the Yamuna River, and is situated on the old waterway bed. The city is additionally celebrated for its wide streets, astounding and current way of life. The below Fig. 2 represents the Map of Greater Noida (See Fig. 3).

More noteworthy Noida is a Greenfield municipality situated in Gautam Budha Nagar locale, Uttar Pradesh. Arrangement in 1991 under the Industrial Area Act of Uttar Pradesh, the municipality is spread over around 90,000 sections of land.

The Authority, in satisfying its duty of making imperative foundation for guaranteeing great personal satisfaction, has received a productive strong waste administration framework in Greater Noida. The framework tends to Solid waste assortment, transportation, treatment (contingent upon the idea of waste produced) and removal. Assortment and capacity of MSW is a significant issue which has incredible effect on the general strong waste administration. Network/metropolitan stockpiling based framework was generally utilized in our urban areas. Recently, there is expanded weight on entryway step assortment for better assistance conveyance and controlling littering just as unapproved dumping of metropolitan trash. The utilitarian component of assortment incorporates not just the get-together of strong squanders and recyclable materials, yet in addition the vehicle of these materials, after assortment, to the area where the assortment vehicle is discharged. This area might be materials handling office, a depot, or a landfill removal site.

Waste Management is the assortment of strong waste from purpose of creation (private, mechanical business, institutional) to the point of treatment or removal. Metropolitan strong waste is gathered in a few different ways. In House-to-House assortment, Waste gatherers visit every individual house to gather trash. The client for the most part pays an expense for this administration. In Community Bins, Users carry their trash to network canisters that are put at fixed focuses in an area or region. MSW is gotten by the region, or its assign, as indicated by a set timetable. In Curbside Pick-Up, Users leave their trash straightforwardly outside their homes as indicated by a trash get plan set with the nearby specialists (optional house-to-house gatherers not normal). While considering oneself conveyed, the generators convey the waste legitimately to removal destinations or depots, or recruit outsider administrators (or the district). In Contracted or Delegated Service, the Businesses employ firms (or district with metropolitan offices) who mastermind assortment timetables and accuses of clients. Regions frequently permit private administrators and may assign assortment zones to energize assortment efficiencies. Source isolation of waste can be demanded and pressurized with inhabitants where there is house-to-house assortment framework is presented. Likewise, it might be accomplished by putting separate canisters for every class of squanders. In any case, the target of isolation could be accomplished when offices are accommodated treatment/preparing of the isolated squanders. As a rule, the isolated burns through are stirred up again at the hour of transportation and removal. Isolation at the assortment point/source diminishes the expense of handling and disposal. Separation at site by cloth pickers for reusing. The practical component of move and transport includes two stages. The exchange of squanders from the littler assortment vehicles, for example, tricycles to the bigger vehicle gear, generally at depot and the ensuing transportation of squanders from depot to the handling or removal site. Transport of waste from assortment focus to definite removal site is significant advance for strong waste administration framework.

3. Results and discussion

From households the waste is gathered by in two different ways, initially, by the administration authorities who gather the waste way to entryway. GNIDA (Greater Noida Industrial Development Authority) have furnished them with the little trucks or they are otherwise called auto travelers in which they gather the waste way to entryway. The quantity of auto travelers working in Greater Noida at present is 14 and there limit is around 150 kg. The quantity of work with every auto travelers is 2(one is driver and other is

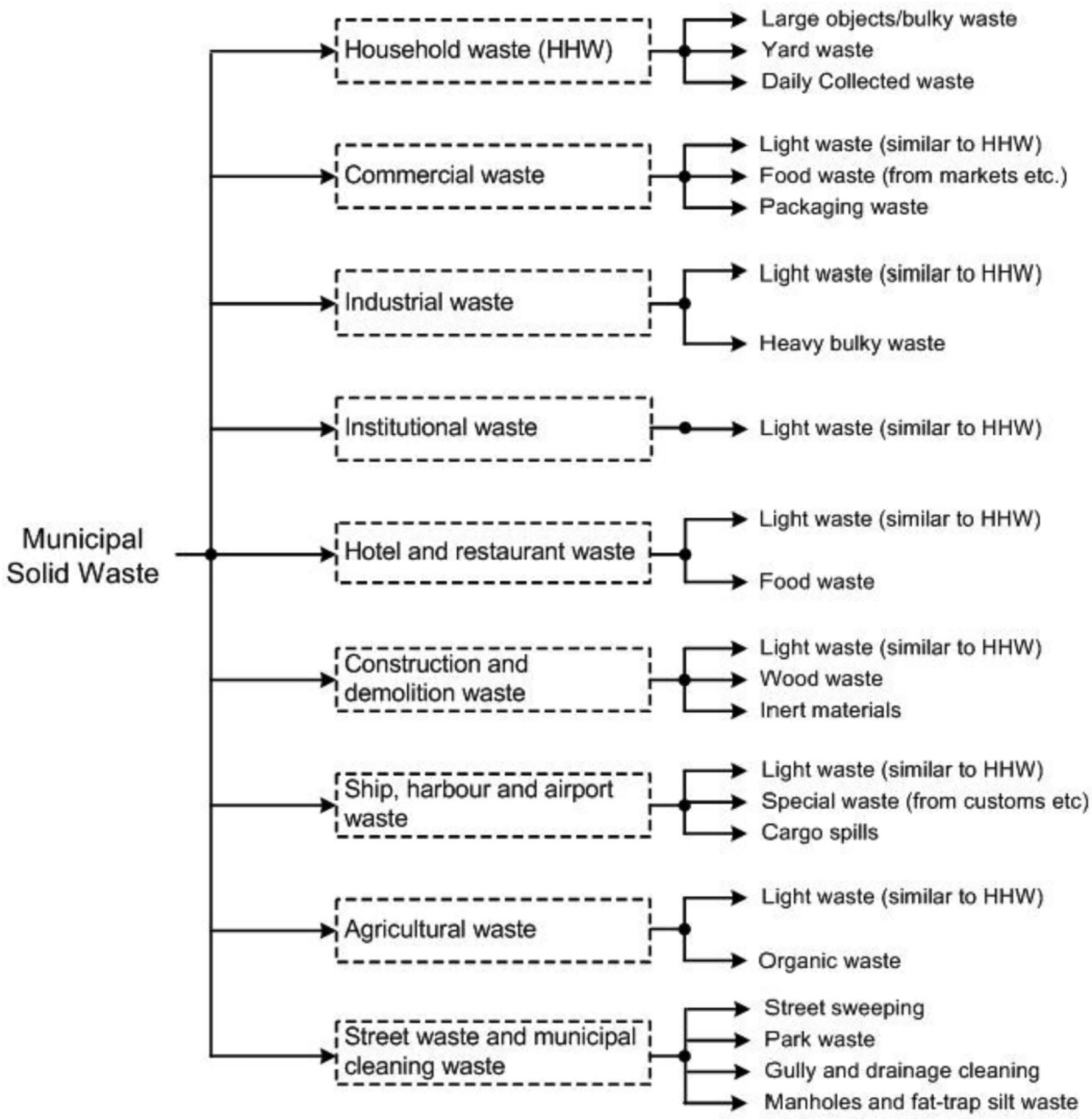


Fig. 1. Municipal Solid Waste composition (Source: IL&FS, MoEF: 32).

collector). Further the gathered waste is taken by the auto sightseers to the large water powered trucks which are put away from the neighborhoods and further these trucks are arranged on dumping site. Furthermore, again by the administration gatherers which gather the waste way to entryway in a cart wherein they store the waste and dump it legitimately on the removal site. The quantity of cart working as per GNIDA (Greater Noida Industrial Development Authority) is 80 in entire territory and there limit is around 120 kg. GNIDA (Greater Noida Industrial Development Authority) has likewise allocated 793 sweepers to clear the waste and gather it in a major size dustbin and which can be additionally gathered by the farm vehicles a short time later.

Transportation of wastes having a major link between the overall gathering and management and removal of wastes at the landfill regions. Mainly for these reason appropriate vehicles are required. The Municipal solid wastes are initially collected using the capacity holder by the network containers and then conveyed to the handling and removal destinations by utilizing a waste collection vehicle. GNIDA (Greater Noida Industrial Development Authority) keeps the process of handling and management of wastes with the help of vehicles also transportation and collection of solid wastes in the municipal surroundings, then getting disposal to waste dumping yards. from. GNIDA (Greater Noida Industrial Development Authority) is utilizing different kinds of trucks also, in the front-End loading machinery also used for vehicle stacking which has to be done in the dustbins. Assessment is done in the ranges from 60% and up to 75% the waste is gathered and gets transported to the waste removal destinations dump yards. As of late, GNIDA (Greater Noida Industrial Development Authority) utilizes various vehicles and loading setup scenario. Total numbers of vehicles are 22 also number of loaders is 45 and four mechanical sweepers. In GNIDA, autotipers are used in that region around is 14. From the houses they are used for the collection of wastes. For transporting the waste to removal site, heavy hydraulic

Fig. 2. Map of Greater Noida.

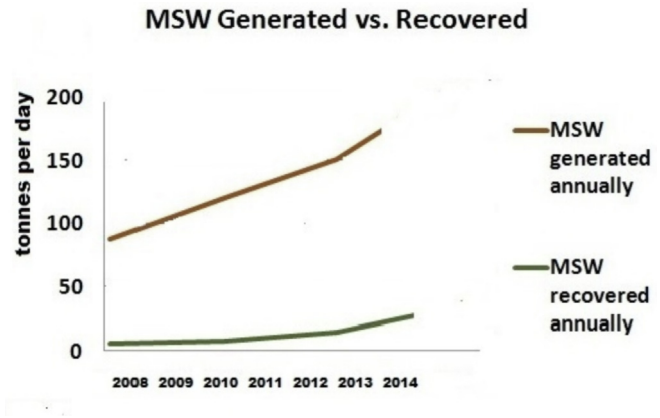


Fig. 3. Municipal Solid Waste Generation Growth (Source: GNIDA).

machineries as well as vehicles are utilized. In the GNIDA zones the volume of accessible vehicles are lacking, the working of vehicles done in two way movements and one excursions was made for every move contingent on the separation of wastes in the removal site. There was delay in moving the waste from the collection of wastes to removal site happens because of utilizing the heavy vehicles. Significant theory and explanations are behind this scenario and it was further advised to properly keep up the vehicles and erratic leaving of vehicles, lost in time and sitting tight on the line for fuel purpose and venturing out starting with one workshop then onto the next for fuelling and so on. Thus, the vehicles are found to work just barely any hours in a day. The circumstance is additionally declined by helpless support and appropriate course arranging and lack of staff. In the region of Greater Noida, two informal sites one was located in Tusyanney and other was in Haldonimorh, Habibpur-Chauhanpur Road, these two arranged for the removal of wastes which are directly are at suit. GNIDA was not liable in the administration of removal waste dumping destinations which currently exist. Zone of removal of site, measure of waste arranged and so forth since current removal site are informal and are by and by contesting. A waste removal site was proposed at Astoli, that comprises a fertilizer plant of 350 tons, in that 250 ton rejected inferred plant.

4. Conclusion

There are various issues in the handling of waste in each and every level, i.e., management of wastes collection, transportation, handling and removal. Batching of Municipal solid waste involves genuine problem for general wellbeing and condition in the surroundings. The conveyance of wastes and transportation vehicles around is lopsided. By and by, there is no landfill destinations in the town for the removal of waste is the most serious issue in Greater Noida. Directly a removal site has been proposed at Astoli in a region of 125 section of land. In which Government must energize treating the soil, vermin-treating the soil, cremation, rejected determined fuel and so forth preparing and treatment techniques for decreasing the strong garbage removal issues in light of the fact that the handling of the waste is just the appropriate response of metropolitan strong waste. A few suggestions would be useful in improving Solid Waste Management (SWM) framework in Greater Noida. Open waste stockpiling locales and other unhygienic road receptacles ought not be permitted. The arrangement of waste containers ought to be right. Isolation of family squander at the source would diminish the weight of strong waste essentially. Appropriate support of vehicles and other equipment's. In overall consideration, the waste management scenario has the major principle of Reduce, Reuse, Recycle and Resource Recovery, but nowadays comparatively 3R's principle is taking place (Reduce, Reuse& Recycle). Government should plan the quantity of landfill and removal site for well support of Municipal Solid Waste (MSW).

CRediT authorship contribution statement

V. Varna Vishakar: Conceptualization, Data curation, Formal analysis, Investigation, Methodology. **B Dhilipkumar:** Validation, Visualization, Writing – original draft. **Ayush Jain:** Supervision, Writing – review & editing. **Zohaib Ahmed Khan:** Project administration, Resources, Software.

Data availability

Data will be made available on request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

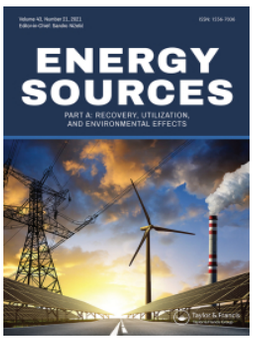
[1] A. Ahsan, Assessment of Municipal Solidwaste management for Developing country, IJRET 1 (4) (2012) 628–641.
[2] "Environmental Hazards of Wastes". www.peta.org / issues / animals – used – for – clothing / leather – industry / leather –environmental – hazards.
[3] S.A. Kumar, Solid waste landfills- Hazardous management, IJST (2004) 3481–3487.
[4] Waste management scenario in Nagpur, National Environment Engineering Research Institute, Nagpur, India, 1996.
[5] J. Singh, Studies on solid waste disposal and management: a review, Ann. NAGI 2 (2004) 74–88.
[6] N. Kumar, Waste Management strategies, Indian J. Environ. Prot. (1999) 90–95.
[7] V. Abubakkar, Waste Management scenario aspects, J. Environ. Manage., Conservat. Recycl. Wastes (2014) 124–145.
[8] S. Vivek, Management of solid wastes in rural and urban areas- Overview, Int. J. Waste Manage. (IJWM) (2008) 36–38.
[9] J. Durai, Municipal Solid waste management of college campus, J. Environ. Protect. (2013) 120–136.
[10] Central Pollution Control Board, Solid Waste Methodology-Overview in Metropolitan cities, 2015.
[11] S. Chakraborty, S.K. Sarkar, Enzyme hydrolysis of solid tannery wastes: Solid-state enzyme production, J. Soc. Leather Technol. Chem. 82 (1998) (1998) 56–58.
[12] M.D. Dianyi Yu, Agency for Toxic Substances and Disease Registry Case Studies in Environmental Medicine (CSEM) Chromium Toxicity http://www.astdr.cdc.gov/csem/csem.asp?csem=10&po=10.
[13] M. Vijaya Sekhar Reddy, K. Ashalatha, K. Sasi, Partial replacement of fine aggregate by tannery shredded waste in concrete, Saudi J. Eng. Technol. 12 (2022).
[14] S. Kumaraguru, S. Tp, C. Rosec, Hydrolysis of tannery fleshings using pancreatic enzymes: Abiotechnological tool for solid waste management, J. Am. Leather Chem. Assoc. 1998 (1998) 32–38.
[15] K. Patchai Murugan, S. Swarnalatha, G. Sekaran, Utilization of Micro Fibrous Carbon Matrix from Tannery Solid Waste for Making Pavement Materials, in: International Conference on Solid Waste Management, 5IconSWM, 2015.
[16] Salman Zafar (January 27,2015). Waste generation in tanneries http://www.bioenergyconsult.com/waste-from-tanneries/.

Varna Vishakar V. currently working as Assistant professor in KIET Group of Institution having Masters in Environmental Specialization from Anna University, Chennai, published 6 International Paper publications in Scopus and other reputed journals. Experienced Chairperson with a demonstrated history of working in the higher education industry. Skilled in Civil Engineering, Environmental Engineering, Environmental Conservation Studies, Impact Assessment, Water Quality Monitoring, Sustainable Development, AutoCAD, Technical Presentation, Project Managing, and Project Coordination.

Dhilipkumar B. is an Assistant Professor in the Department of Civil Engineering at K. Ramakrishnan College of Technology, Samayapuram, Trichy – 621112. Currently doing part-time research at NIT Patna, Ashok Rajpath, Bihar – 800005.

Ayush Jain, currently working as Assistant professor in KIET Group of Institution has masters in Alternate Hydro Energy System from Indian institute of technology Roorkee and completed his graduation in Civil Engineering from Rajasthan Technical University published 2 papers related to the Land use and land cover dynamics.

Zohaib Ahmed Khan is an Assistant Professor in the Department of Civil Engineering at KIET Group of Institutions, Delhi-NCR, Ghaziabad, Completed M.Tech in Hydraulics and Water Resources Engineering from Delhi Technological University.



Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ueso20

Analysis of organic rankine cycle powered by N-number of solar collectors in series

Deepak Tiwari, Deepali Atheaya, Anil Kumar & Nishant Kumar

To cite this article: Deepak Tiwari, Deepali Atheaya, Anil Kumar & Nishant Kumar (2022) Analysis of organic rankine cycle powered by N-number of solar collectors in series, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44:3, 6678-6697, DOI: 10.1080/15567036.2022.2096726

To link to this article: https://doi.org/10.1080/15567036.2022.2096726

Published online: 25 Jul 2022.

Submit your article to this journal

Article views: 33

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ueso20

ENERGY SOURCES, PART A: RECOVERY, UTILIZATION, AND ENVIRONMENTAL EFFECTS
2022, VOL. 44, NO. 3, 6678–6697
https://doi.org/10.1080/15567036.2022.2096726

Taylor & Francis
Taylor & Francis Group

Check for updates

Analysis of organic rankine cycle powered by N-number of solar collectors in series

Deepak Tiwari[a], Deepali Atheaya[b], Anil Kumar[c,d], and Nishant Kumar[e]

[a]Department of Mechanical Engineering, Vivekananda School of Engineering & Technology, New Delhi, India; [b]Department of Mechanical Engineering, Bennett University, Greater Noida, India; [c]Department of Mechanical Engineering, Delhi Technological University, Delhi, India; [d]Centre for Energy and Environment, Delhi Technological University, Delhi, India; [e]Department of Embedded System, National Institute of Electronics & Information Technology, Calicut, India

ABSTRACT

Thermodynamic model of the solar power organic Rankine cycle (ORC) has been proposed in the present study. Model is developed in MATLAB 2018a for Cyclohexane/R123 mixture-based ORC system driven by series of N-collectors. Four different collector models namely: Case-I, "N series connection of photovoltaic thermal conventional compound parabolic (NPVTCCPC) collector system"; "Case-II: N series connection of photovoltaic thermal (NPVT) collector system"; "Case-III: N series connection of conventional compound parabolic collector (NCCPC) system"; Case-IV N series connection of flat plate collector (NFPC) system" have been integrated with ORC system. Overall thermal performance has been evaluated for each case and compared individually. Case-III (NCCPC) emerged as the best combination in the view of highest collector exit temperature 110°C, heat gain 255 kW, thermal efficiency 14.14%, exergetic efficiency 54.77%. Further, highest percentage of total exergy destruction was observed in evaporator (37–45%) followed by condenser (28–31%), expander (24–30%) and lowest in pump (1–2%). The present developed system could be a promising solution to power crises in remote and off-grid-connected areas.

ARTICLE HISTORY
Received 3 May 2022
Revised 20 June 2022
Accepted 24 June 2022

KEYWORDS
N-series of collector; organic Rankine cycle; Cyclohexane/R123; energy; exergy

Introduction

Along with rising income, the amount of carbon dioxide (CO_2) emission in the world has been increased significantly over the years. Due to speedy urbanization and industrialization the CO_2 emission is prevailing in developing countries like China 6.19% afterward India 5.98%, Indonesia 1.29%, and Brazil 1.25% (Alam et al. 2016). Scientists have stressed environmental concerns caused by CO_2 emissions such as global warming, climate change, increased sea level, disturbance in rainfall patterns, and land fertility reduction. Therefore, environmental issues have been forced to consider alternate/renewable and clean energy sources for power generation. Application of renewable energy sources for electricity generation in India can mitigate the escalating energy demand and CO_2 emission by 4 to 45% (Kaur and Luthra 2018). India's last five years, the contributions of renewable energy sources in power generation were only 5 to 9%. However, it was targeted to achieve 175 GW of electricity production from renewable energy sources, out of which 100 GW from solar energy only by year 2022 (Central Electricity Authority Ministry of Power Government of India, 2018–19). India has ample potential to produce electricity by utilizing solar energy because of more than 300 days available with a clear sky along with 1700 to 1900 kWh of available solar radiation (Kumar, Prakash, and Dube 2017). Electricity from a low-grade (less than 350°C) solar heat source can be generated from the combination of photovoltaic and thermal collector integrated with Organic

CONTACT Deepak Tiwari deepakmathiya@gmail.com; deepak.tiwari@vips.edu Department of Mechanical Engineering, Vivekananda School of Engineering & Technology, Outer Ring Rd, AU Block, New Delhi 110034, India

© 2022 Taylor & Francis Group, LLC

Rankine cycle (ORC) (Cioccolanti et al. 2018). Further, the ORC system can also be applied for variable refrigerant flow VFR run by evacuated thermal collector for combined heating and cooling (Gilani et al. 2022).

Overall, collector exit temperature relies on beam radiation, mass flow rate and temperature of the working fluid, concentration ratio, incident modifier angle, transmittance for incident, tilt angle, conduction, and convection heat losses (Atheaya et al. 2016a). Apart from this, collector exit temperature can also be enhanced by adding more collectors in series. However, same outlet water temperature appears at collectors no. 6 and 11 when increasing 6 to 11 in series for photovoltaic thermal compounded parabolic concentrator (PVTCPC) collectors (Tripathi and Tiwari 2016). Moreover, the collector (PVTCPC) is thermodynamic efficient and cost effective (Aheaya, Tiwari, and Tiwari 2016b).

The novelty of the present study is that several researchers have evaluated the ORC system's thermodynamic performance integrated with different kinds of only single concentrator. It includes glazed reversed absorber compound parabolic concentrator (Tiwari et al. 2017) Glazed reversed absorber PVT-CPC (Tiwari et al. 2020), solar pond (Zeynali, Akbari, and Khalilian 2019), heliostat (Habibi et al. 2020), parabolic concentrator (Cioccolanti, Hamedani, and Villarini 2019) evacuated flat plate collector (Kutlu et al. 2019), parabolic trough collector (Elakhdar et al. 2019) and flat plate collector (Wang et al. 2013). It is summarized in Table 1. As per the literature review, performance of solar power organic Rankine cycle coupled with different configurations of N-number of collectors has not been carried out so far in any of the research article (Table 1). Therefore, the objective of the present work is to perform thermodynamic modeling for ORC system driven by four different configurations of N collectors engaged in series, namely; N-photovoltaic and thermal conventional compound parabolic concentrator (NPVTCCPC, Case-I), N-photovoltaic and thermal collector (NPVT, Case-II,), N-conventional compound parabolic concentrator (NCCPC, Case-III,) and N flat plate collector (NFPC, Case-IV).

In this manuscript we have taken six collectors in series (N = 6) for all cases (Case-I, Case-II, Case-III and Case-IV). The mass flow rate is fixed (0.012 kg/ sec) and the N collectors (6) are of series connection to obtain the fluid outlet temperature which varies with respect to time. However, if we will fix the temperature say (90°C) then the effect of number of collectors can be investigated. Researchers have indicated that analogous outlet water temperature is achieved at lower collectors for PVTCPC owing to solar concentration radiation (Tripathi and Tiwari 2016).

Manuscript has the following sections: introduction, description of the proposed system, selection of the potential working fluid, implementation of mass and energy conservation, and balance of exergy in thermodynamic analysis for combined solar ORC system. Another portion of the work covers a detailed description of results and concludes the outcomes of the present study and its scope for further extension.

Table 1. Review of solar powered ORC.

Author	Collector	Working fluid	Remark	Reference
Tiwari et al. (2020)	Reverse absorber PVT-CPC	Heptane/R245fa	Maximum collector outlet temperature, heat gain and thermal efficiency at 1:00 hrs with six concentration ratios	(Tiwari et al. 2020)
Zeynali et al. (2019)	Solar pond	R152a, R123, R600, R245fa, R134a	The regenerative cycle with open feed water and internal heat exchanger exhibits best among basic ORC, ORC with IHE and ORC with OFLH	(Zeynali, Akbari, and Khalilian 2019)
Habibi et al. (2020)	Heliostat	CO_2, CO, Oxygen, Nitrogen, Air and Helium	Helium exhibits the highest work output and exergetic efficiency	(Habibi et al. 2020)
Cioccolanti et al. (2019)	Parabolic compound concentrator	R245fa	Solar multiple is the most affecting parameter and increasing area of the collector reduces the environmental impact	(Cioccolanti, Hamedani, and Villarini 2019)
Elakhdar et al. (2019)	Parabolic concentrator	R601a, R245fa, R141b and R123	Increasing evaporator temperature COP enhances and entraining ratio increases	(Elakhdar et al. 2019)
Wang et al. (2013)	Flat plate collector	R245fa, R123, Isobutane, R134a	R245fa and R123 concluded as potential working fluid at low pressure	(Wang et al. 2013)

Description of the proposed model

Solar heat source based organic Rankine cycle power system installed with various configurations of collectors in N-series is proposed. The description of proposed model is described as follows.

Simple ORC system

A simple organic Rankine cycle system consists of four main components: evaporator, expander, condenser, and pump. Overall structure of the organic Rankine cycle resembles with conventional Rankine cycle system; however, smaller in size. Organic Rankine cycle also exhibits relatively better cycle performance and lowest operational and maintenance costs among various low-grade power-generating cycles such as the Trilateral, Goswami, and Kalina.

Figure 1 shows schematic of the solar ORC system integrated with collectors in series. Working fluid in the evaporator (collector) gets heated and converted into vapor from sun's beam radiation (shown by process 2–4, Figure 2). Saturated fluid enters the expander (Turbine), where power is generated after the enthalpy drop (shown by process 4–5, Figure 2). Superheated vapor exit from the expander gets converted into a saturated liquid through the condenser (shown by process 5–1, Figure 2). Saturated liquid enters the evaporator from the condenser through the pump (shown by process 1–2, Figure 2), where the pressure of the working fluid increases; thereby, the cycle completes. Figure 3 shows the cross-sectional side view of the N-series of collectors. Outlet temperature of the N^{th} collector increases due to N-number of collectors installation in series.

Exit of the first collector is connected with the inlet of the second collector and exit of the second collector connects with the inlet of the third collector. This way N-number of collectors connects with the series. Working principle of four different N-series connected collectors, namely Case-I, Case-II, Case-III, and Case-IV, have been described are as follows:

Case-I: N series connection of photovoltaic thermal conventional compound parabolic (NPVTCCPC) collector system

Figure 4 shows the series-connected NPVTCCPC. Solar radiation falls to the reflector plate of the compound parabolic concentrator system. It further comes to the metal absorber, which is partially covered with photovoltaic cells and glass.

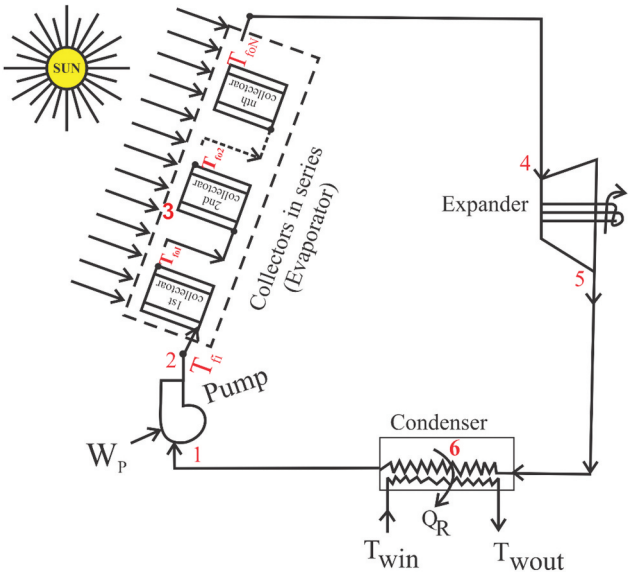


Figure 1. Simple organic Rankine cycle (ORC).

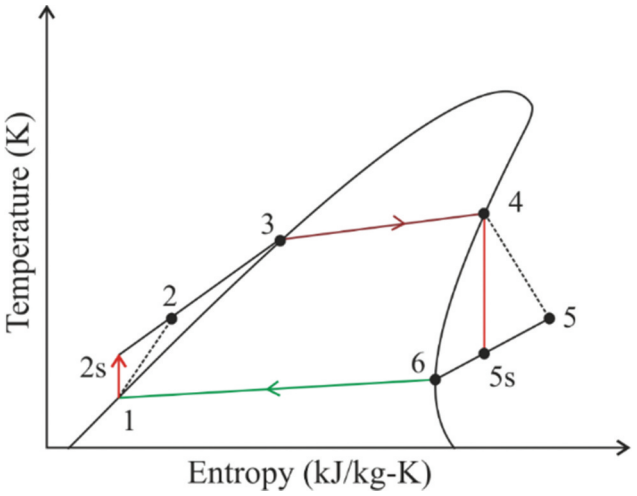


Figure 2. Temperature-entropy diagram for basic structure of ORC employing mixture working fluid.

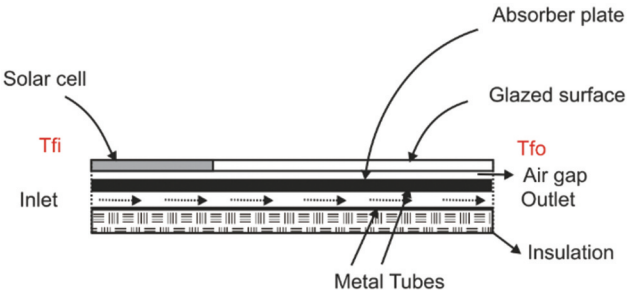


Figure 3. Sectional view of the N-series collector.

Thermal energy gets transferred to the water flowing through the water pipes by conduction and convection heat transfer process from the backside of the solar cells and glass, thereby rise in water temperature. PVT-CCPC of N-systems are of series connection to enhance the outlet water temperature at the N^{th} collector system.

Case-II: N series connection of photovoltaic thermal (NPVT) collector system

Figure 5 exhibits the schematic diagram of series-connected NPVT. Here compound parabolic concentrator is detached from the system (Case-I). Solar radiation falls on the photovoltaic thermal collector system and the thermal energy obtained from the back of solar cells and glass has been utilized to increase thermal energy. Further, partially covered photovoltaic thermal collectors are of series connection to increase the temperature of the water.

Case-III: N series connection of conventional compound parabolic collector (NCCPC) system

Figure 6 represents the schematic diagram of series integrated NCCPC (Case-III). Here, the photovoltaic cells are detached from Case-I. The solar radiation comes to the reflector plate and gets reflected to the absorber plate of CCPC system to impart thermal energy. Further, compound parabolic concentrator systems are of series connection, giving rise to higher water temperature flowing in the pipes below the absorber plate.

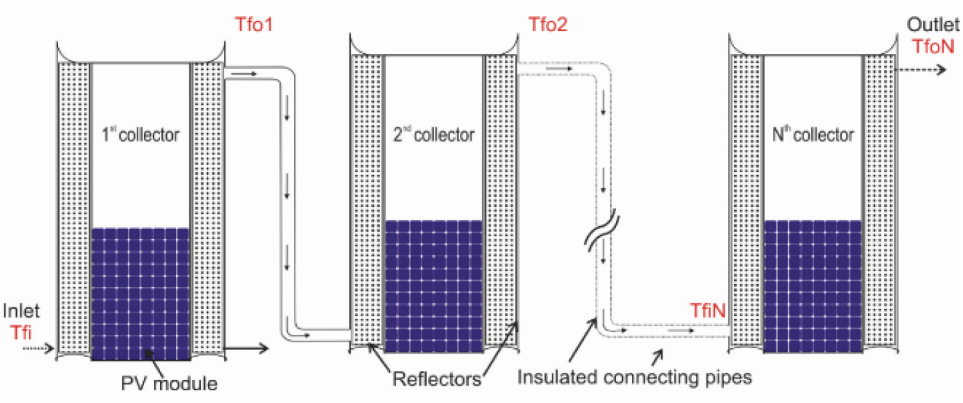


Figure 4. N series connection of photovoltaic thermal conventional compound parabolic (NPVTCCPC) collector system.

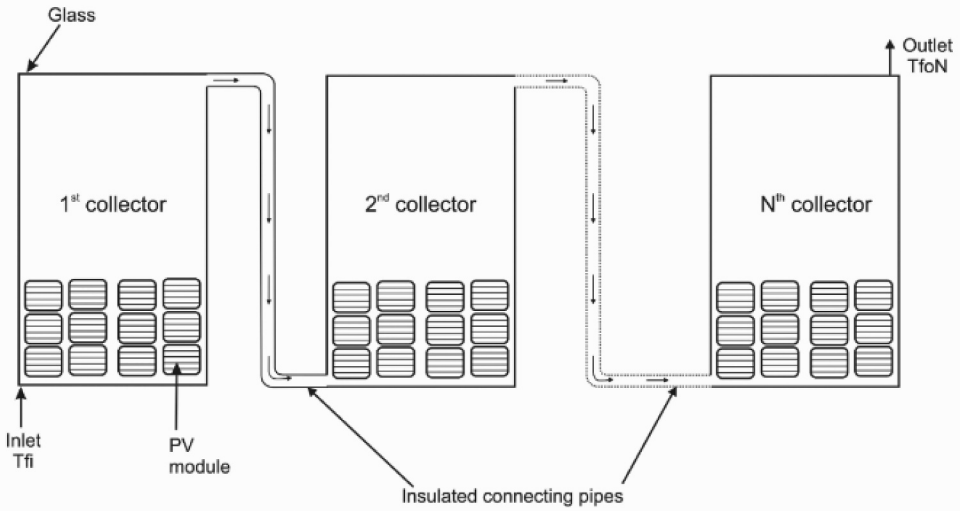


Figure 5. N-series photovoltaic thermal collector (N-PVT).

Case-IV: N series connection of flat plate collector (NFPC) system

Figure 7 exhibits the schematic diagram of series-connected NFPC (Case-IV). The photovoltaic cells and compound parabolic concentrator system are absent. Solar radiation directly falls to the flat plate collector system to get thermal energy. Flat plate collector system is assembled with a series connection in order to increase the exit temperature.

Selection of working fluid and thermodynamic model

Mixture working fluid felicitate the better heat match between source fluid to sink in heat exchanger that improves the thermodynamic performance of ORC than pure fluid.

Cyclohexane/R123 has opted as a working fluid at a definite mass proportion of 0.1/0.9. Mass proportion 0.1 has been taken for flammable working fluid cyclohexane, whereas the comparatively more mass proportion of 0.9 of nonflammable R123 was chosen to make a nonflammable working fluid mixture (Table 2). Moreover, cyclohexane shows 11 global warming potential, whereas R123 exhibits 77. Ozone

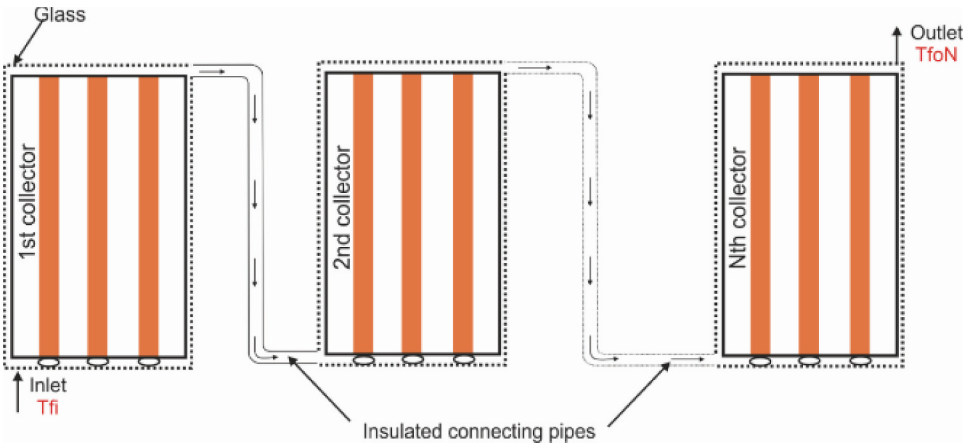


Figure 6. N-series conventional compound parabolic concentrator.

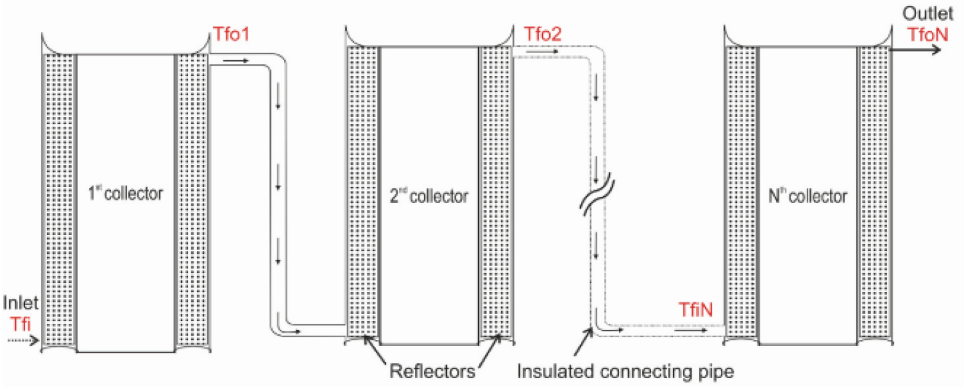


Figure 7. N-series flat plate collector.

depletion potential of cyclohexane is zero. The working fluid R123 shows a negligible ozone depletion potential of 0.020. Therefore, working fluid cyclohexane/R123 is safe in use and does lesser harm to the environment.

Thermodynamic modeling is done based on conservation of mass, energy, and exergy balance (Moran and Shapiro 2010). Following assumptions were considered during thermodynamic modeling of the proposed system:

- All solar power ORC system components are treated as control volume system.
- Solar power ORC is in steady-state mode. Changes in kinetic and potential energy have been neglected.
- Heat losses in heat exchanger devices and their connecting pipes are neglected.
- Pressure drops are neglected for all in evaporator, condenser and also in associated pipes (Shu et al. 2014).
- The water (incompressible) considered for working fluid in collector.

Table 2. Thermal characteristic of working fluid (Tripathi and Tiwari 2016).

Working Fluid	Boiling Temperature (K)	Critical Temperature (K)	Critical Pressure (MPa)	Nature of Fluid	Flammability	ODP	GWP	Toxic
R-245fa	288.29	427.01	3.651	Dry	No	0	950	Yes
Cyclohexane	353.89	553.64	4.075	Dry	Yes	0	Low	Non-toxic
R123	300.97	467.7	3.66	Wet	No	0.020	77	Toxic

Conservation of mass

Mass flow rate of the working fluid in one inlet and exit point for the system is given as:

$$\left(\frac{dm}{dt}\right)_{cv} = \sum_{i} \dot{m}_i - \sum_{e} \dot{m}_e \tag{1}$$

where 'cv' denotes the control volume, $'\dot{m}'$ denotes the mass flow rate of the working fluid, 'i' denotes the inlet and 'e' shows the exit.

Energy balance

Energy balance rate for a single inlet and exit system is expressed by:

$$\left(\frac{dE}{dt}\right)_{cv} = \dot{Q}_{cv} - \dot{W}_{cv} + \sum_{i} \dot{m}_i \left(h_i + \frac{V_i^2}{2} + gz_i\right) - \sum_{e} \dot{m}_e \left(h_e + \frac{V_e^2}{2} + gz_e\right) \tag{2}$$

where 'E' denotes the total energy of the system, '$\dot{Q}$' shows the rate of heat transfer, $\dot{W}$ exhibits the rate of work transfer and 'h' exhibits the enthalpy, $'V'$ shows stream flow velocity, $'g'$ exhibits acceleration due to gravity and $'z'$ denotes the elevation from the datum.

Exergy balance

Exergy analysis is an actual method to appraise the merit of energy conversion or distribution processes and systems. It is the process by which a source of inefficiency of any thermal system can be determined and overcome. Exergy flow for single-stream (single inlet and exit) of the system is expressed by:

$$\left(\frac{dE_{ex}}{dt}\right)_{cv} = \sum_{k} \left(1 - \frac{T_o}{T_k}\right) \dot{Q}_k - \dot{W}_{cv} + p_o \frac{dv_{cv}}{dt} + \sum_{i} \dot{m}_i ex_{fi} - \sum_{e} \dot{m}_e ex_{fe} - \dot{E}D \tag{3}$$

The term $'E_{ex}'$ exhibits the total exergy of any system, $'\sum_{k}\left(1 - \frac{T_o}{T_j}\right)\dot{Q}_j$ shows the exergy transfer rate in terms of heat at the 'k' location of the boundary and $'\dot{W}_{cv}'$ rate of exergy transfer in terms of work whereas $'\frac{dv_{cv}}{dt}'$ represents the rate of change of volume and $'p_o'$ shows the pressure at the ambient.

Energy and exergy analysis

The exit temperature, balance of thermal energy for solar collectors are analyzed as follows (Singh and Tiwari 2017):

Solar collector system

(i) Temperature at the exit of the NPVTCCP Collector (Case-I) is computed by the expression:

$$T_{foN} = \left[\left\{\frac{(1-K_k^{(N)})}{(1-K_k)}\right\}\right] \times \frac{1}{\dot{m}_{wf,c}C_{wf,c}}\left[I_b \times (AF_R(\tau\alpha))_1 + T_a \times (AF_R U_L)_1\right] + T_{fi}K_k^{(N)} \quad (4)$$

Expression is used to compute the exit temperature of the Nth collector for photovoltaic thermal (NPVT) water collector system of series connection, Case-II:

'$I_t(t) = I_b(t)$'; '$A_{am} = A_{rm}$'; '$A_{rm} = A_m$']

$$T_{foN} = \left[\left\{\frac{(1-K_k^{'(N)})}{(1-K_k^{'})}\right\}\right] \times \frac{1}{\dot{m}_{wf,c}C_{wf,c}}\left[I_t \times (AF_R(\tau\alpha))_1^{'} + T_a \times (AF_R U_L)_1^{'}\right] + T_{fi}K_k^{'(N)} \quad (5)$$

Temperature at the exit of the NCCCP Collector (Case-III) is computed by the expression:

('$A_{am} = A_{rm}$'; '$A_m = A_{rm} = 0$')

$$T_{foN} = \left[\left\{\frac{(1-K_k^{''(N)})}{(1-K_k^{''})}\right\}\right] \times \frac{1}{\dot{m}_{wf,c}C_{wf,c}}\left[I_b \times (AF_R(\tau\alpha))_1^{''} + T_a \times (AF_R U_L)_1^{''}\right] + T_{fi}K_k^{''(N)} \quad (6)$$

(iv) Expression is used to compute the exit temperature of the Nth collector when N number of flat plate water collector (NFPC) system of series connection, Case-IV: ('$A_m = 0$,' '$A_r = A_a$')

$$T_{foN} = \left[\left\{\frac{(1-K_k^{'''(N)})}{(1-K_k^{'''})}\right\}\right] \times \frac{1}{\dot{m}_{wf,c}C_{wf,c}}\left[I_t \times (AF_R(\tau\alpha))_1^{'''} + T_a \times (AF_R U_L)_1^{'''}\right] + T_{fi}K_k^{'''(N)} \quad (7)$$

(v) Overall thermal gain for each case can be computed by the expression:

$$Q_{ov,th,eg,N} = \sum Q_{th,eg,N} + \sum \frac{Q_{el,eg,N}}{\gamma} \quad (8)$$

$'\gamma = 0.38'$ denotes the factor of conversion from thermal energy to electrical energy in coal power plant.

(vi) The gain in thermal energy for each case of considered collectors:

$$Q_{th,eg,N} = \dot{m}_{wf}c_{p,c} \times (T_{foN} - T_{fi}) \quad (9)$$

(vii) Equation (10) is used to evaluate electrical energy gain for Case-I, Case-II (Shyam et al. 2015):

$$Q_{el,eg,N} = I_b \times A_m \sum_1^N \eta_{m,N} \quad (10)$$

(viii) Equation (11) computes the photovoltaic module efficiency for (Case-1, Case-II:

$$\eta_{m,N} = \tau_g \beta \eta_{c,N} \quad (11)$$

(ix) Equation (12) evaluates the overall efficiency of the collector for Case-I (NPVTCCPC) and Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC):

$$\eta_{ov,th,eff,c} = \frac{Q_{ov,th,eg,N}}{N(A_c + A_m)I_t} \quad (12)$$

(x) Equation (13) shows the exergy contained with working fluid in the collector:

$$\dot{EX}_{ov,EX,NC} = \sum \dot{EX}_{th,NC} + \sum \dot{EX}_{el,NC} \quad (13)$$

(xi) Thermal exergy is evaluated by Equation (14) (Bejan and Porvaznik 1978):

$$\dot{EX}_{th,NC} = \dot{m}_{wf,c} \times C_{p,wf,c}(T_{foN} - T_{fi}) - \dot{m}_{wf} \times C_{pwf,c}(T_a + 273)\ln\left(\frac{T_{fo}+273}{T_{fi}+273}\right) \tag{14}$$

(xii) Equation (15) evaluates the exergy of the solar photovoltaic (SPV) module:

$$\dot{EX}_{el} = \rho \times I_b(t) \times \eta_{el,ex,m} \times A_m \tag{15}$$

where $'\eta_{el,ex,m}'$ denotes the electrical exergetic efficiency for module.

(xiii) Input exergy delivered by beam radiation (Petela 2003):

$$EX_i = I_b(t)\left\{1 - \frac{4}{3}\left(\frac{T_o}{T_s}\right) + \frac{1}{3}\left(\frac{T_o}{T_s}\right)^4\right\} \tag{16}$$

where $'T_a'$ exhibit the ambient temperature in 'K' and $'T_s = 600'$ shows the temperature of the sun in 'K'.

(xiv) Overall collector exergetic efficiency is evaluated by the Equation (17) (Szargut 2003):

$$\eta_{ov,EX,NC} = \eta_{m,N} + \frac{\dot{EX}_{th,NC}}{A_{NC} \times EX_i} \tag{17}$$

Energy and exergy analysis of organic Rankine cycle (ORC) system

(i) Power consumed by the pump is defined by the Equation (18):

$$W_{p_{-1,2}} = \dot{m}_{wf} \times (h_2 - h_1) \tag{18}$$

(ii) Equation (19) evaluates the heat generated in the evaporator (collector):

$$Q_{g_{-2,4}} = \dot{m}_{wf} \times (h_4 - h_2) \tag{19}$$

(iii) Equation (20) computes the work deliver by expander shaft:

$$W_{t_4,5} = \dot{m}_{wf} \times (h_4 - h_5) \tag{20}$$

(iv) Heat rejected to the ambient from condenser is evaluated by equation (21)

$$Q_{cond_5,1} = \dot{m}_{wf} \times (h_5 - h_1) \tag{21}$$

(v) Thermal efficiency is evaluated by the Equation (21):

$$\eta_{th_ORC} = \left(\frac{W_{t_4,5} - W_{p_{-1,2}}}{Q_{g_{-2,4}}}\right) \tag{22}$$

(vi) Using Equation (22) the exergetic efficiency of the ORC system is evaluated:

$$\eta_{EX_orc} = \left(\frac{W_{t_4,5}}{EX_{i,orc}}\right) \tag{23}$$

(vii) Input of exergy to the ORC system computes the Equation (23)(Cengel and Boles 2011):

$$EX_{i,orc} = \left\{Q_{g_{-2,4}} \times \left(1 - \frac{T_o}{T_{hs}}\right) + W_{p_{-1,2}}\right\} \tag{24}$$

(viii) Equation (24) calculates the total exergy destruction in ORC system:

$$\dot{EX}_{D,total} = (\dot{EX}_{D,\mathrm{p}} + \dot{EX}_{D,\mathrm{ev}} + \dot{EX}_{D,\mathrm{exp}} + \dot{EX}_{D,\mathrm{cond}}) \tag{25}$$

(ix) Destruction of exergy in pump is defined by the Equation (25):

$$\dot{EX}_{D,p} = \dot{m}_{wf} \times T_o(s_2 - s_1) \tag{26}$$

(x) Equation (26) calculates the destruction of exergy in evaporator:

$$\dot{EX}_{D,\mathrm{ev}} = \dot{m}_{wf} \times T_o\left\{(s_4 - s_2) - \left(\frac{h_4 - h_2}{T_{hs}}\right)\right\} \tag{27}$$

(xi) Destruction that happens in expander estimated by the Equation (27):

$$\dot{EX}_{D,\mathrm{exp}} = \dot{m}_{wf} \times T_o(s_5 - s_4) \tag{28}$$

(xii) Destruction of exergy destruction that took place in condenser is expressed by Equation (28):

$$\dot{EX}_{D,\mathrm{cond}} = \dot{m}_{wf} \times T_o\left\{(s_5 - s_1) - \left(\frac{h_5 - h_1}{T_o}\right)\right\} \tag{29}$$

(xiii) Altogether efficiency for thermal and exergetic for solar power ORC system expressed by Equation (29) and (30) (Wang et al. 2010):

$$\eta_{ov,th_sorc} = \eta_{th_ORC} \times \eta_{ov,\mathrm{th},eff,c} \tag{30}$$

And in the same way exegetic efficiency can also be calculated as:

$$\eta_{ov_EX_sorc_NC} = \eta_{EX_orc} \times \eta_{ov,\mathrm{EX},NC} \tag{31}$$

Validation

Table 3 expresses the perfect match of the collector output in terms of heat gain for collectors in Case-I (NPVTCCPC), Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC) and also have been validated with (Tripathi and Tiwari 2016). Moreover, percentage deviation is encountered 0.16–0.48% for present study and (Wu, Zhu, and Yu 2016). Therefore, it validates the findings of ORC system (Table 4).

Result and discussion

Mathematical modeling of the solar ORC system integrated with various N-collectors engaged in series is done in MATLAB 2018a. Working mixture (fluid) properties received from REFPROP 9.0 (Lemon, Huber, and Mclinden 2010) interlinking with MATLAB 2018a. Properties (thermophysical) of the opted zeotropic mixture 'Cyclohexane/R123' working fluid derived by REFPROP. The present communication studied the collector exit temperature, heat gain in the evaporator (collector), collective efficiency for thermal and exergetic efficiency as per the hours (08:00 – 16:00 hrs) and further evaluated exergy loss. Input parameters considered in mathematical thermal modeling for various collector systems are given in Table 5. Table 6 lists the values of input parameters used for ORC system modeling.

Hourly (08:00 – 16:00 hrs) global radiation, ambient temperature, and beam radiation of Delhi (India) for clear sky day for January have been taken from the Indian meteorological department (IMD) Pune (India). Hourly variation in beam radiation and global radiation is shown in Figure 8. Maximum values of global radiations (733.9 W/m^2) and beam radiation (597.2 W/m^2) were observed at 13.00 hrs. Thermodynamic modeling predicted performance parameters of solar ORC system coupling with the various collector of Case-I (NPVTCCPC), Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC) are described as:

Table 3. Validation of solar collectors.

	Thermal energy gain (W) Case-I		Thermal energy gain (W) Case-II		Thermal energy gain (W) Case-III		Thermal energy gain (W) Case-IV	
Time(hrs)	Present work	(Tripathi and Tiwari 2016)	Present work	(Tripathi and Tiwari 2016)	Present work	(Tripathi and Tiwari 2016)	Present work	(Tripathi and Tiwari 2016)
08:00	0.55	0.98	0.7	0.7	0.8	1.59	0.7	0.61
09:00	1.35	1.49	1.75	1.00	2.5	3.98	1.75	0.94
10:00	2.3	3.89	2.5	1.87	3.8	4.68	2.7	3.87
11:00	2.75	3.98	3	4.76	4.75	5.63	3.2	4.06
12:00	2.95	2.81	3.2	5.2	5	4.05	3.35	4.86
13:00	2.95	3.80	3.25	2.88	5.05	3.08	3.45	2.69
14:00	2.75	3.00	3	2.04	4.5	3.67	3.15	2.56
15:00	2.25	3.32	2.45	3.79	3.7	2.69	2.5	1.98
16:00	1.5	1.86	1.75	0.98	2.5	1.68	1.75	0.90

Table 4. Validation of ORC system.

		Thermal efficiency (%)		
Working fluids	Mass fractions	Present study	(Wu, Zhu, and Yu 2016)	Deviation (%)
R245fa /Butane	0.9/0.1	8.94	8.99	0.55
R245fa/R227ea	0.9/0.1	8.95	9.07	0.88
RC 318/R245fa	0.9/0.1	8.95	9.07	0.45

Table 5. Design parameters and constants used in solar thermal collector (Mc Adams 1954).

Parameter	Value	Parameter	Value
$A_{r,m}$	0.25 m^2	k_g	0.816 $W/m\,K$
A_r	1 m^2	k_p	6 $W/m\,K$
A_{rc}	0.75 m^2	L_g	0.003 m
A_{ac}	1.5 m^2	L_i	0.100 m
A_{am}	0.5 m^2	A_a	2 m^2
C_f	4.179 k $J/kg\,K$	L_p	0.002 m
F_{rm}	0.8110 m^2	$\dot{m}_f$	0.012 kg/s
k_i	0.166 $W/m\,K$	PF_1	0.3782
PF_2	0.9512	PF_c	0.9842
h_{pf}	100 $W/m^2\,K$	h_i	$h_i = 5.7 + 3.8V \quad 0 \le V \le 7$
h'_i	5.8 $W/m^2\,K$	h_o	$V = 0, h_i = 5.7\,W/m^2K$ $h_o = 5.7 + 3.8V\ W/m^2K$ $V = 1m/s, h_o = 9.5\ W/m^2K$
F'	0.9680	F_{rc}	0.8693 m^2
$U_{l,c}$	4.7 $W/m^2\,K$	U_{L1}	3.47 W/m^2K
$U_{t,pa}$	4.8 $W/m^2\,K$	$U_{t,ca}$	9.17 W/m^2K
$U_{t,cp}$	5.58 $W/m^2\,K$	τ_g	0.95
η_o	0.15	α_p	0.8
α_c	0.9	ρ	0.84
β_c	0.89	F_{rc}	0.86 m^2

Table 6. Design input parameter for ORC (Tiwari et al. 2020).

Input parameters	Values
Evaporation Temperature at bubble state	Variable
Mass fraction ratio for mixture working fluid	0.1/0.9
Isentropic efficiency for expander	0.80
Isentropic efficiency for pump	0.70
Condensation temperature at dew state	35°C

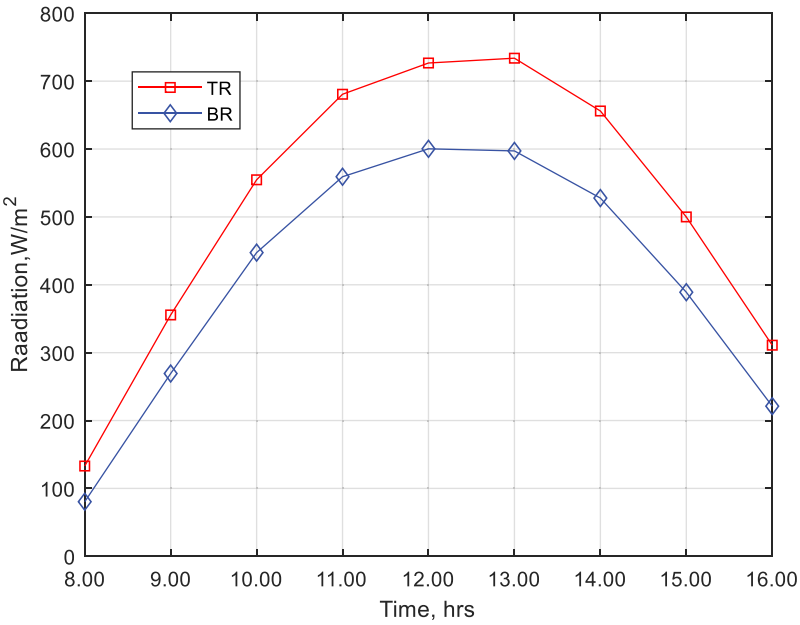


Figure 8. Hourly beam and total radiation (Ref. Indian Metrological Department Pune, India).

Heat gain in collectors

Hourly heat gain generated in the collector (evaporator) was calculated for Case-I (NPVTCCPC) Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC) as shown in Figure 9.

Maximum heat gain 2.55×10^5 W was achieved in Case-III followed by 2.44×10^5 W in Case-I, then 2.43×10^5 W in Case-II and least 2.34×10^5 in Case-IV at 13:00 hrs. Beam radiation enters through aperture area, gets transferred from the reflector, and eventually reaches the absorber. It absorbs in case of only conventional compound parabolic concentrator for Case-III, resulting in maximum thermal energy (heat) gain. Further, in Case-I (NPVTCCPC), absorber area is covered by half of PV module and the remaining half by metal plate. Therefore, relatively less heat absorbed into the absorber plate relatively less leads to lower thermal energy gain than Case-III. Heat gain in Case-II is lower than Case-I because total absorber surface is covered with PV module only, due to which heat gain received at the absorber surface is lower than Case-I. Least heat gain in Case-IV is caused by notable heat losses by radiation, convection and conduction due to greater collector area.

Inlet and exit temperature for various collector

Figure 10 exhibits the variation of fluid inlet and exit temperature for N-number of collectors of series connection in different cases, Case-I (NPVTCCPC) and Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC) for hourly basis (8:00 hrs-16:00 hrs).

Fluid inlet temperature for all collectors is a maximum 18°C at 16:00 hrs and minimum 8.4°C at 12:00 hrs. Highest exit fluid temperature 110°C found in Case-III, 92°C in Case-I, 90°C in Case-II and least 76°C for Case-IV at 13:00 hrs. The extent of temperature rise in collector outlet temperature strongly depends on the amount of thermal energy generated (gain) at the absorber surface, which is transferred to the tube's working fluid. Collector exit temperature for Case-I is higher than Case-II and lower than Case-III because it shows a higher heat gain than Case-II and lower heat gain than Case-III.

Highest collector output temperature obtained in Case-III because it contains highest thermal energy gain among Case-I, Case-II and Case-IV.

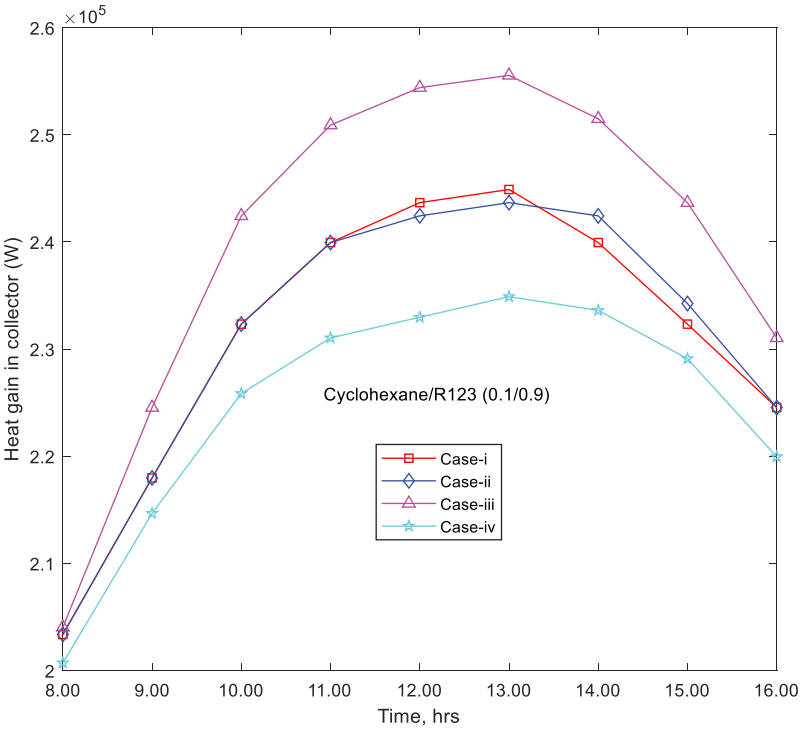


Figure 9. Hourly heat gain in various collector.

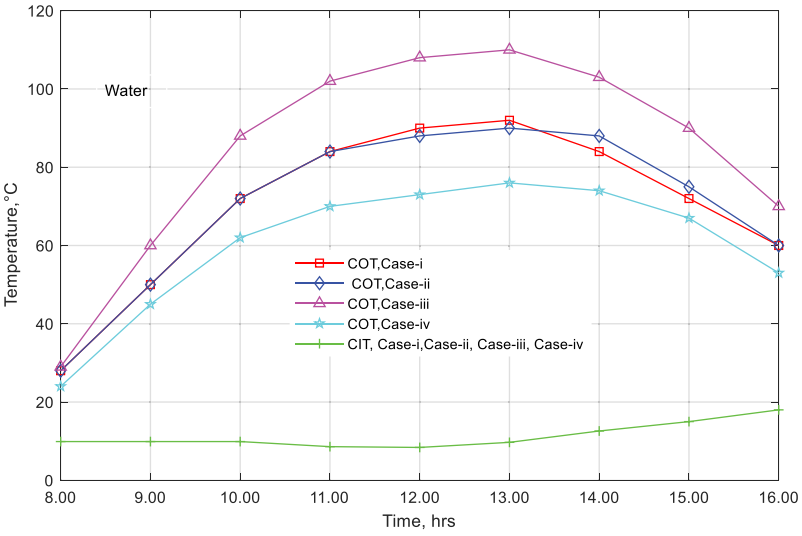


Figure 10. Hourly outlet and inlet temperature in various collector.

Overall thermal and exergetic efficiency

Hourly thermodynamic performance for solar power ORC system for all cases is depicted in Figure 11. Combined thermodynamic efficiency is 14.14% for Case-III, 12.2% in Case-I, 11.96% Case-II and least 10.12% in Case-IV at 13:00 hrs. Overall thermal efficiency of solar power ORC system depends upon the extent of heat generated (gain) in the collector and resulting exit temperature. Thermal efficiency of combined solar power ORC increases when collector exit temperature increases.

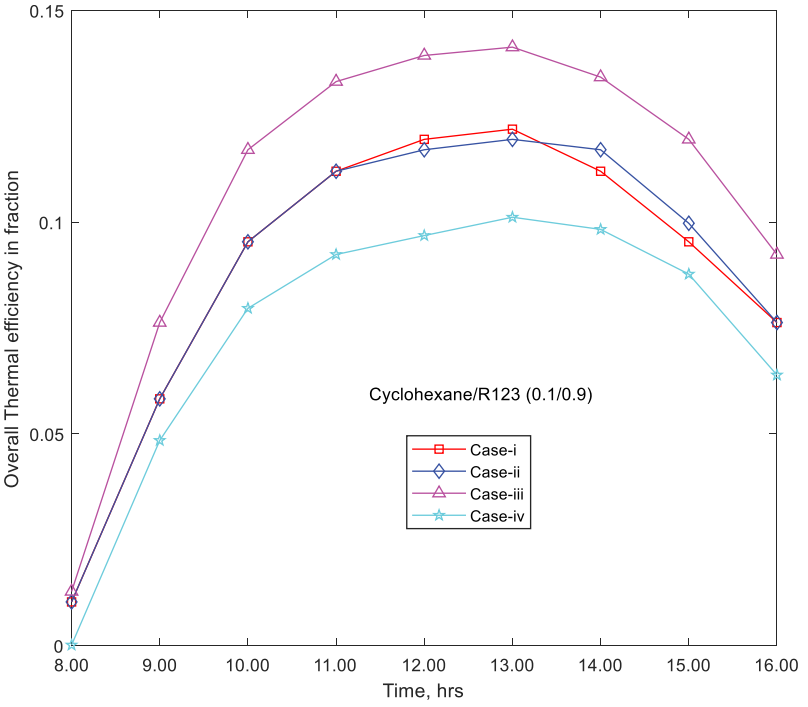


Figure 11. Hourly thermal efficiency.

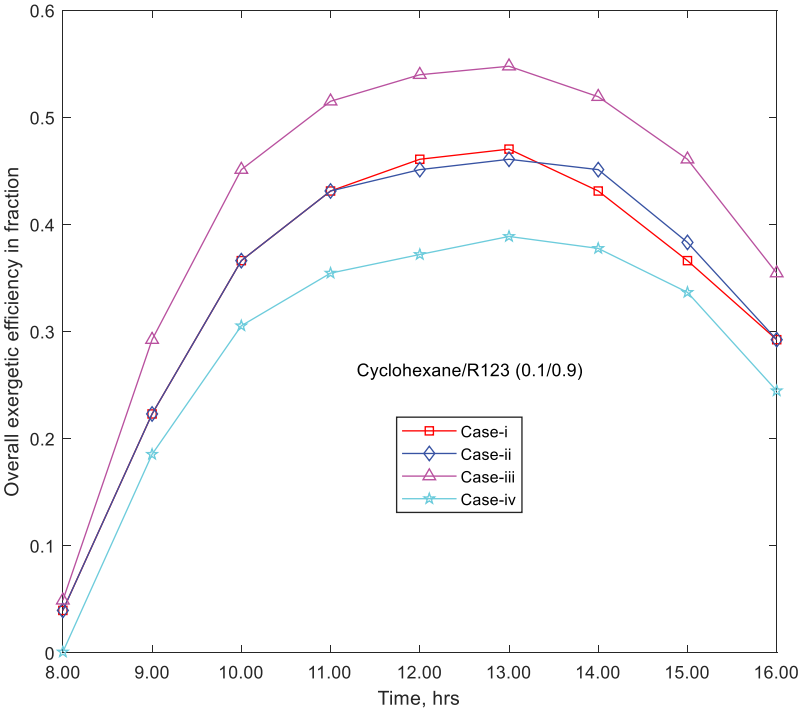
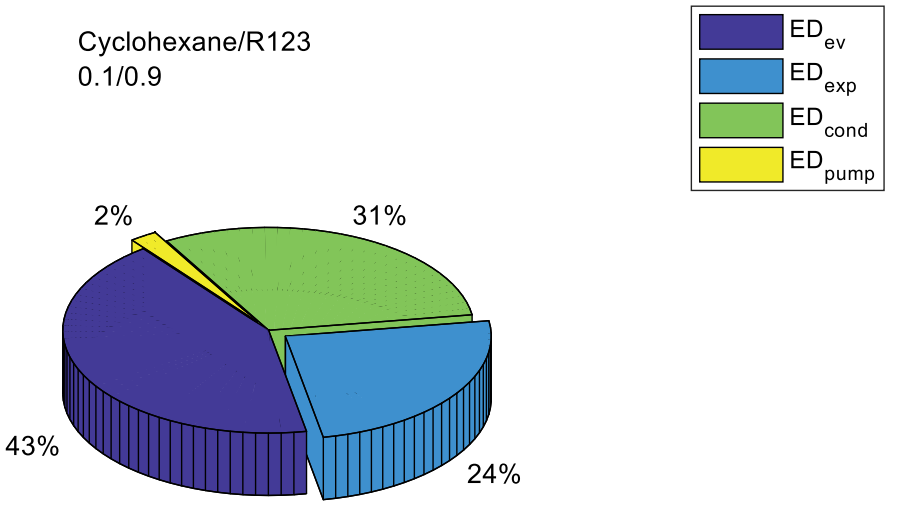
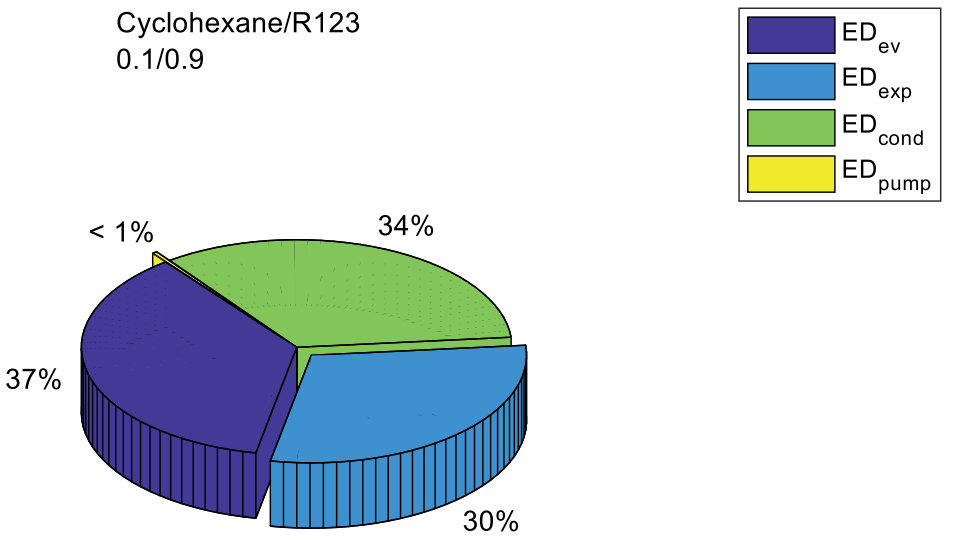


Figure 12. Hourly exergetic efficiency.

The overall thermal efficiency is lowest in Case-IV caused by lowest thermal energy gain and resulting lowest collector exit temperature. The same trend appears in exergetic efficiency for Case-I (NPVTCCPC), Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC), as shown in Figure 12. Highest overall exergetic efficiency is achieved, 54.77% in Case-III and the lowest 38.87% in Case-IV, whereas Case-I exhibits 47.03% and Case-II delivers only 46.08% at 13:00 hrs.



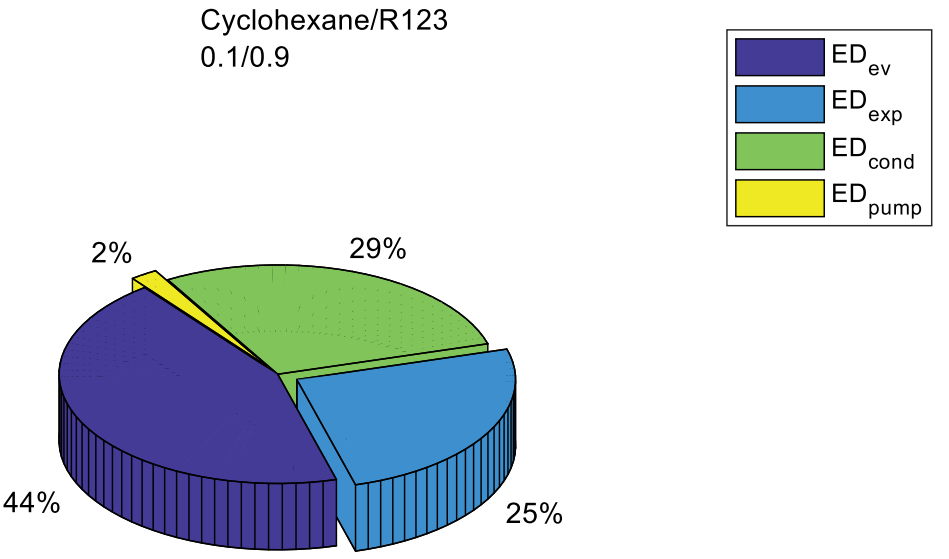
a DestructionofExergy in ORC system parts for Case-I



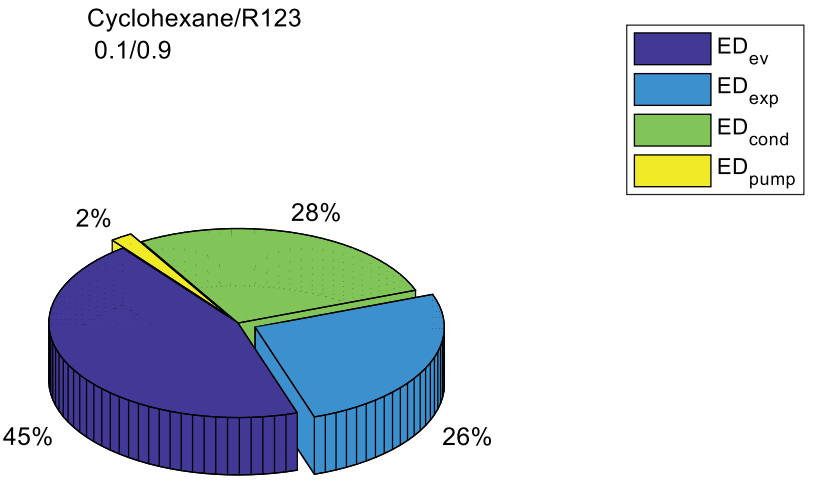
b Destruction of Exergy in ORC system parts for Case-II

Figure 13. (a) Destruction of Exergy in ORC system parts for Case-I. (b) Destruction of Exergy in ORC system parts for Case-II. (c) Destruction of exergy in ORC system parts Case-III. (d) Destruction of exergy in ORC system parts for Case-IV.

The exergetic efficiency in Case-III is highest owing to increased heat gain and expander output. Exergetic efficiency in Case-I is inferior to that of the Case-III due to lower enthalpy drop and heat gain than Case-III. Moreover, exergetic efficiency of Case-I is higher than Case-II and case-IV because of higher heat gain and enthalpy drop than Case-II and Case-IV. Exergetic efficiency of combined solar power ORC



c Destruction of exergyin ORC system partsCase-III



d Destruction of exergyin ORC system partsfor Case-IV

Figure 13. (Continued)

depends on extent of heat gain in collector and resulting expander output. When exit temperature of the collector increases which enhance the heat gain in collector as well as enthalpy drop in expander. Therefore Case-IV shows the lowest exergetic efficiency among Case-I, Case-II and Case-III at 13:00 hrs.

Exergy destruction

Figure 13a-d reflects exergy destruction in various component for ORC system while powered by Case-I, Case-II, Case-III and Case-IV respectively. Further, major portion exergy destruction was contributed by evaporator (37–45%) then after condenser (28–31%), expander (24–30%) and least in pump (1–2%) for all cases at 13:00 hrs for constant mass proportion (0.1/0.9) of the working fluid cyclohexane/R123. The evaporator and condenser contribute major exergy loss owing to varying temperature difference between hot (source) and cold fluid (sink).

Conclusions

In this study, series connection of N collectors of different configurations coupled with ORC system has been thermodynamically compared on an hourly basis. Major outcomes that have emerged are as follows:

- Case-III (NCCPC) exhibits the highest temperature 110°C followed by 92°C in Case-I (NPVTCCPC), 90°C for Case-II (NPVT) and least 76°C for NFPC (Case-IV).
- Highest thermal energy gain exhibited by NCCPC collectors (Case-III) 2.55×10^5 W, then 2.44×10^5 W for Case-I (NPVTCCPC) after that 2.43×10^5 W in Case-II (NPVT) and lowest 2.34×10^5 W in Case-IV (NFPC).
- Overall thermal efficiency 14.14% of Case-III (NCCPC) collector highest among 12.2% in Case-I (NPVTCCPC), 11.96% for Case-II (NPVT) and least 10.12% in Case-IV (NFPC).
- Overall exergetic efficiency 54.77% emerged as highest in the set of Case-III (NCCPC) collectors, followed by 47.03% in Case-I (NPVTCCPC), Case-II (NPVT) 46.08% and Case-IV (NFPC) 38.07%.
- In all four configurations of ORC system driven by Case-I (NPVTCCPC) and Case-II (NPVT), Case-III (NCCPC) and Case-IV (NFPC), the major percentage of exergy destruction contributed by evaporator (37–45%) after that condenser (28–31%) next after expander (24–30%) and least by pump (1–2%).

The present study should extend to fabricate the solar power ORC system. ORC-based power generation systems can be designed on a large scale for remote areas where power distribution and transmission losses are significant. Moreover, for establishing such a system driven by renewable heat sources like solar, many subsidies are given to investors in developing countries like India, China, Brazil, Indonesia, etc.

Acknowledgments

Authors express our heartfelt thanks to Vivekananda School of Engineering and Technology (VSE&T), Delhi, India for accommodating data processing facility to accomplish this research article.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Deepali Atheaya http://orcid.org/0000-0002-2693-0295

Anil Kumar http://orcid.org/0000-0002-2846-4347

Nomenclature

Abbreviation	
A	*Area* (m^2)
A_m	*Area of module* (m^2)
A_c	*Area of the collector* (m^2)
$A_{r,m}$	*Receiver area covered with module*(m^2)
A_r	*Receiver area* (m^2)
A_{rc}	*Receiver area covered through glass*(m^2)
A_{ac}	*Aperture area covered with glass* (m^2)
A_{am}	*Module covered aperture area* (m^2)
BR	*Beam radiation* (W/m^2)
$C_{p,wf}$	*Specific heat of the working fluid* ($J/kg\,K$)
CIT	*Collector inlet temperature* ^{o} C
COT	*Collector outlet temperature* ^{o} C
GWP	*Global warming potential*
h	*Specific enthalpy (kJ/kg)*
h_{pf}	*Convective heat transfer coefficient between plate and water* (W/m^2K)
I_t	*Total radiation* (W/m^2)
I_b	*Beam radiation* (W/m^2)
$\dot{m}$	*Mass flow rate of the working fluid in (kg/s)*
ORC	*Organic Rankine cycle*
ODP	*Ozone depletion potential*
$\dot{Q}$	*Heat flow rate (W)*
s	*Specific entropy (kJ/kg K)*
T	*Temperature (K)*
TR	*Total radiation* (W/m^2)
U_L	*Overall heat transfer coefficient* (W/m^2K)
$\dot{W}$	*Work transfer rate (W)*
Greek letters	
$(\alpha\tau)_c$	*Transmissivity of the collector and product of effective absorptivity and reflectivity*
τ	*Transmittivity*
α	*Absorptivity*
η	*Efficiency*

Subscripts	
a	*Aperture*
b	*Beam*
$cond$	*Condenser*
c	*Collector*
cv	*Control volume*
eg	*Energy gain*
eff	*Efficiency*
el	*Electrical*
ov	*Overall*
o	*Ambient*
EX	*Exergy*
ex	*Exergy*
exp	*Expander*
ev	*Evaporator*
$\dot{E}D$	*Exergy destruction rate (W)*
e	*Exit*
i	*Inlet*
fe	*Fluid exit*
fi	*Fluid inlet*
m	*Module*
NC	*N-collectors*
th	*Thermal*

References

Alam, M. M., M. W. Murad, M. A. H. Ozturk, and I. Ozturk. 2016. Relationships among carbon emissions, economic growth, energy consumption and population growth: Testing environmental kuznets curve hypothesis for Brazil, China, India and Indonesia. *Ecological Indicators* 70:466–79. doi:10.1016/j.ecolind.2016.06.043.

Atheaya, D., A. Tiwari, G. N. Tiwari, and I. M. Al-Helal. 2016a. Performance evaluation of inverted absorber photovoltaic thermal compound parabolic concentrator (PVT-CPC): Constant flow rate mode. *Applied Energy* 167:70–79. doi:10.1016/j.apenergy.2016.01.023.

Atheaya, D., A. Tiwari, and G. N. Tiwari. 2016b. Exergy analysis of photovoltaic thermal (PVT) compound parabolic concentrator (CPC) for constant collection temperature mode. *Solar Energy* 135:222–34. doi:10.1016/j.solener.2016.05.055.

Bejan, A., and M. J. Porvaznik. 1978. General criterion for rating heat-exchanger performance. *Heat Mass Transfer* 21 (2):655–58. doi:10.1128/iai.21.2.655-658.1978.

Cengel, Y. A., and M. A. Boles. 2011. *Thermodynamics: An engineering approach*. 7th ed. New York: Tata McGraw Hill.

Cioccolanti, L., R. Tascioni, E. Bocci, and M. Villarini. 2018. Parametric analysis of a solar ORGANIC Rankine cycle trigeneration system for residential applications. *Energy Conversion and Management* 163:407–19. doi:10.1016/j.enconman.2018.02.043.

Cioccolanti, L., S. Hamedani, and M. Villarini. 2019. Environmental and energy assessment of a small-scale solar organic Rankine cycle tri-generation system based on compound parabolic collectors. *Energy Conversion and Management* 198:111829. doi:10.1016/j.enconman.2019.111829.

Elakhdar, M., H. Landoulsi, B. Tashtoush, E. Nehdi, and L. Kairouani. 2019. A combined thermal system of ejector refrigeration and organic Rankine cycles for power generation using a solar parabolic trough. *Energy Conversion and Management* 199:111947. doi:10.1016/j.enconman.2019.111947.

Gilani, H. A., S. Hoseinzadeh, F. Esmaeilion, S. Memon, D. A. Garcia, and M. E. H. Assad. 2022. A solar thermal driven ORC-VFR system employed in subtropical Mediterranean climatic building. *Energy* 250:123819. doi:10.1016/j.energy.2022.123819.

Habibi, H., M. Zoghi, A. Chitsaz, K. Javaherdeh, M. Ayazpour, and E. Bellos. 2020. Working fluid selection for regenerative supercritical Brayton cycle combined with bottoming ORC driven by molten salt solar power tower using energy–exergy analysis. *Sustainable Energy Technologies and Assessments* 39:100699. doi:10.1016/j.seta.2020.100699.

Kaur, R. R., and A. Luthra. 2018. Population growth, urbanization and electricity - Challenges and initiatives in the state of Punjab, India. *Energy Strategy Review* 21:50–61. doi:10.1016/j.esr.2018.04.005.

Kumar, A., O. Prakash, and A. Dube. 2017. A review on progress of concentrated solar power in India. *Renewable and Sustainable Energy Reviews* 79:304–07. doi:10.1016/j.rser.2017.05.086.

Kutlu, C., M. T. Tahir Erdinc, J. Li, Y. Wang, and Y. Su. 2019. A study on heat storage sizing and flow control for a domestic scale solar-powered organic Rankine cycle-vapour compression refrigeration system. *Renewable Energy* 143:301–12. Accessed 2019 May 17. renene. doi:10.1016/j.renene.2019.05.017.

Lemon, E. W., M. L. Huber, and M. Mclinden. 2010. NIST standard reference database 23: Reference fluid thermodynamic and transport properties–REFPROP.9.0.

Mc Adams, W. A. 1954. *Heat Transmission*, 249. 3rd ed. New York: Mc Graw-Hill Book Co.

Moran, M. J., and H. N. Shapiro. 2010. *Fundamental of engineering thermodynamics*. 6th ed. UK: John Willey & Sons Ltd.

Petela, R. 2003. Exergy of undiluted thermal radiation. *Solar Energy* 74 (6):469–88. doi:10.1016/S0038-092X(03)00226-3.

Shu, G., Y. Gao, H. Tian, H. Wei, and X. Liang. 2014. Study of mixtures based on hydrocarbons used in ORC (organic Rankine cycle) for engine waste heat recovery. *Energy* 74:428–428. doi:10.1016/j.energy.2014.07.007.

Shyam, T. G. N., Tiwari, G. N. Al-Helal, and I. M. Al-Helal. 2015. Analytical expression of temperature dependent electrical efficiency of NPVT water collectors of series connection. *Solar Energy* 114:61–76. doi:10.1016/j.solener.2015.01.026.

Singh, D. B., and G. N. Tiwari. 2017. Performance analysis of basin type solar stills integrated with N identical photovoltaic thermal (PVT) compound parabolic concentrator (CPC) collectors: A comparative study. *Solar Energy* 142:144–58. doi:10.1016/j.solener.2016.11.047.

Szargut, J. T. 2003. Anthropogenic and natural exergy losses (exergy balance of the Earth's surface and atmosphere). *Energy* 28 (11):1047–54. doi:10.1016/S0360-5442(03)00089-6.

Tiwari, D., A. F. Sherwani, D. Atheaya, and A. Arora. 2017. Energy and exergy analysis of solar driven recuperated organic Rankine cycle using glazed reverse absorber conventional compound parabolic concentrator (GRACCPC) system. *Solar Energy* 155:1431–42. doi:10.1016/j.solener.2017.08.001.

Tiwari, D., A. F. Sherwani, D. Aheaya, A. Kumar, and N. Kumar. 2020. Thermodynamic analysis of organic Rankine cycle driven by reversed absorber hybrid photovoltaic thermal compound parabolic concentrator system. *Renewable Energy* 147:2118–27. doi:10.1016/j.renene.2019.10.018.

Tripathi, R., and G. N. Tiwari. 2016. Energetic and exergetic analysis of N partially covered photovoltaic thermal-compound parabolic concentrator (PVT-CPC) collectors of series connection. *Solar Energy* 137:441–51. doi:10.1016/j.solener.2016.08.048.

Wang, X. D., L. Zhao, J. L. Wang, W. Z. Zhang, X. Z. Zhao, and W. Wu. 2010. Performance evaluation of low-temperature solar Rankine cycle system utilizing R245fa. *Solar Energy* 84 (3):353–64. doi:10.1016/j.solener.2009.11.004.

Wang, M., J. Wang, Y. Zhao, P. Zhao, and Y. Dai. 2013. Thermodynamic analysis and optimization of solar driven regenerative organic Rankine cycle (ORC) based on flat-plate solar collectors. *Applied Thermal Engineering* 50 (1):816–25. doi:10.1016/j.applthermaleng.2012.08.013.

Wu, Y., Y. Zhu, and L. Yu. 2016. Thermal and economic performance analysis of zeotropic mixture for organic Rankine cycles. *Applied Thermal Engineering* 96:57–63. doi:10.1016/j.applthermaleng.2015.11.083.

Zeynali, A., A. Akbari, and M. Khalilian. 2019. Investigation of the performance of modified organic Rankine cycles (ORCs) and modified trilateral flash cycles (TFCs) assisted by a solar pond. *Solar Energy* 182:361–81. doi:10.1016/j.solener.2019.03.001.

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/0957-4093.htm

Being ethically resilient during COVID-19: a cross-sectional study of Indian supply chain companies

Shivangi Verma and Naval Garg
University School of Management and Entrepreneurship, Delhi Technological University, Delhi, India, and
Thangaraja Arumugam
Business School, Vellore Institute of Technology, Vellore, India

Received 2 May 2022
Revised 14 June 2022
21 June 2022
5 July 2022
Accepted 6 July 2022

Abstract

Purpose – The present study aims to examine the relationship between techno-ethical orientation and ethical decision-making (EDM) in Indian supply chain companies during the COVID-19 pandemic. It also aims to explore the moderating role of technological frames (TF) in the relationship between techno-ethical orientation and EDM.

Design/methodology/approach – The relationship between techno-ethical orientation and EDM is examined using correlation and regression analysis. The moderating effect of five dimensions of TFs (personal attitude, application value, organisational influence, supervisor influence and industry influence) is analysed using structural equation modelling.

Findings – The correlation coefficient between techno-ethical orientation and EDM is 0.513. Also, the regression coefficient ($\beta = 0.213$) is significant at 0.05, establishing a positive linkage between the two. *R*-square values showed a 45.2% variation in EDM is explained by techno-ethical orientation. Similarly, all variables of TFs have a positive and significant moderating effect on the relationship between techno-ethical orientation and EDM.

Originality/value – This is one of the pioneer studies exploring techno-ethical orientation's impact on EDM in supply chain companies.

Keywords Techno-ethics, Ethical decision-making, Technological frames, India

Paper type Research paper

1. Introduction

These days, most of the research, studies and surveys revolve around the impact of coronavirus in almost all sectors, in one way or another. Although the effects of the pandemic are not uniformly spread throughout the industries and countries, it is important to note that there exists disparity in managing the aftermath of COVID-19. While some countries strategically and judiciously allocated their resources to their best and optimum utilisation to combat the impact of COVID-19, other countries have, to some extent, lagged to bounce back effectively and survive in such adversity. The COVID-19 pandemic has caused widespread hardships for various industries across the globe. The state-imposed lockdowns, shortage of raw material, import-export disruptions, fewer product development and poor consumer demand severely impacted the survival and growth of modern organisations. The pandemic has severely halted the functioning of significant industries, thereby affecting the economic and financial development of the country (Khan *et al.*, 2022). During the pandemic, the supply chain industry was among the most pertinent and profoundly impacted sectors (Harapko, 2021; Sarkis, 2021; Sriyanto *et al.*, 2022). Due to the global supply chain disruption, demand has decreased; production has slowed; inventory cost has increased; and the vendor-supplier relationship has been severely hampered (McKibbin and Fernando, 2020; PwC, 2021). Such instability influences supply chain operations, making them vulnerable to several disruptions (Dubey *et al.*, 2021; Ivanov, 2020; Shishodia *et al.*, 2021). Guan *et al.* (2020)

The International Journal of Logistics Management
© Emerald Publishing Limited
0957-4093
DOI 10.1108/IJLM-05-2022-0203

reported that the lockdown scenario during coronavirus alone cost the global supply chain a loss of US$ 20.0tn under a 2-month period to US$ 30.1tn (equivalent to 40.3% of global value-added) under a 6-month period.

Though low-income and middle-income countries were highly vulnerable to indirect effects, the countries less affected by the virus also experienced huge losses. It highlights the importance of risk mitigation and building a resilient supply chain to avoid massive shortages of materials and delays in delivery during unpredicted events. Such disruptions cause ripple effects and depreciate the performance and value of the supply chain (Shishodia *et al.*, 2021). Recently, studies are emphasising the call for action to work with uncertainty and cope with future similar recurrences through learning from adversarial events (Calia *et al.*, 2021; Rajesh, 2022; Shishodia *et al.*, 2021). Responding quickly, being prepared for disruptions, and adapting to changes can help develop a resilient network that can withstand shortfalls and reduce vulnerabilities (Caschili *et al.*, 2015; Ho *et al.*, 2015; Ivanov *et al.*, 2017; Scholten and Schilder, 2015). India, a developing and emerging market, was at the centre of attention before and after the pandemic due to its commendable performance in the supply chain industry. Before the pandemic, India was regarded as an important factor in the global supply chain. Even during the pandemic, the country effectively repaired its damaged value chains in the short term. Despite the challenges raised by COVID-19, the nation was ranked second in the 2021 Agility Emerging Markets Logistics Index (IBEF, 2021). This success is attributed to its vast distribution network, resilient supply chain and technological advancement.

Supply chain companies adopted different alternatives and strategies to restructure their operations and build a resilient, cost-effective and sustainable supply chain (Sarkar *et al.*, 2022). Companies have accelerated their supply chain interactions with the help of digitisation and the introduction of new technologies to drive effective decision-making in the complex and challenging process (Eshkenazi, 2018; Laaper *et al.*, 2018). These digital initiatives include electronic customer service management, blockchain, cloud computing, enterprise planning system and augmenting online order procurement and placement (Asnordin *et al.*, 2021; Sanders and Morgan, 2020). Such novel technologies and analytical methods in supply chain management aided in building competitive advantage and enabled the supply chain operations to boost (Buntz, 2017; Newman, 2018; Vollmer, 2018). The introduction of information technology (IT) paved the way for logistics 4.0 and the digital supply chain. These initiatives in supply chain industries enhanced organisational capacity for sustainable thinking (Kumar *et al.*, 2020). The adoption of blockchain technologies, data-driven systems, Internet of Things (IoT) and automation has internalised digital decision-making by supply chain managers (van Hoek, 2019; Saberi *et al.*, 2019). The technology during COVID-19 allowed industries to develop resilience and survivability that can withstand large-scale disruptions (Sarkar *et al.*, 2022).

With the increased use of information and communication technology (ICT) at the workplace, the organisation receive considerable benefits in terms of productivity, low operating costs, improving product quality, generating new business development opportunities, flexibility, timely solutions and increased communication (Chatterjee *et al.*, 2021; Demeester *et al.*, 2021; Parungao, 2020). Nonetheless, concerns regarding the ethical use of technology in the workplace are rising. Although digitalisation in supply chain companies has led to revolutionary improvements in the design and manufacturing processes, operations and services, it has also increased ethical complexities (Tjahjono *et al.*, 2017). Some of the technology-related ethical issues experienced by organisations during the recent global health pandemic were a cyber-security risk, data tracking, data stealing, Internet spammers perpetrating attacks, inadequate bandwidth of the unencrypted and unregulated channel, unpatched software, obsolete operating systems, inability to ensure safety and confidentiality of sensitive data (Verma and Garg, 2022; Parungao, 2020).

Employees and organisations working in the virtual environment and connected via technology have reported unethical usage of modern technologies and ICT devices (Carroll and Conboy, 2020; Verma and Garg, 2022). Moreover, owing to the widespread adverse effects of coronavirus and the mandatory adoption of technology for communication, the unethical business practices during the COVID-19 period have been exacerbated. Given the pervasive and inevitable use of technology, researchers have started to develop antecedents, forms and aftermath of unethical use. Several authors have demonstrated that unethical business practices increase supply chain costs, result in dissatisfied customers and lower customer service levels (Simangunsong *et al.*, 2016). The key technology-related unethical practices in the supply chain industry that emerged or undergone tremendous remodification due to the uncertain outbreak of the pandemic are grouped under the following heads (Bag *et al.*, 2018; López-Fernández and Sánchez-Rosales, 2022; Yun *et al.*, 2020):

(1) Individual-level – The introduction of digital supply chain and cloud computing has questioned the use of data for business operations. Data piracy, infringement of confidentiality and data tracking are rising. Employees use available data for cyber-attacks, bullying, threatening vendors/suppliers, misleading customers/clients and distorting the value chain for their benefit. Stealing the data about vendors, suppliers, customers and designs to create their setup or defame colleagues/supervisors/organisations on various social media platforms poses a severe threat to the integrity of supply chain companies.

(2) Group level – The recent surge in demand for toilet papers and hand sanitisers or masks are examples of suppliers forming cartels and artificially creating scarcity by withholding the supply and inflating the price. This activity ensures "panic buying" among customers, which is usually controlled by a few suppliers of a local market. Copying the product information and making a similar product using technology, especially 3D printing, is another concern for escalating unethical behaviour.

(3) National level – Lack of government regulations, policies and management is one of the many reasons accelerating the misuse of technology by a particular industry. Though supply chain companies utilise blockchain technology, big data analytics and cloud computing for their smooth operations, their regulation and policy for transparent utilisation and implementation are under development. Meanwhile, companies and individuals take undue advantage of the opportunity and gaps in the ecosystem.

(4) Global level – With technology in place, the whereabouts of shipments, their origins and destinations, contents, locations, etc., are all tracked across global delivery networks. This raises the question of data tracking, which may exploit and disrupt the value chain of any country.

Considering the widespread misuse of resources and technology at various levels, studying techno-ethics in the supply chain is paramount. Larger unethical practices and scandals may lead to even more catastrophic consequences for the firms when the use of technology is indispensable for business operations. The impact of unethical conduct sometimes reaches far beyond the firm's control, such as a significant drop in the firm's valuation and stock prices, creating spillover effects, distrust between supply chain partners, customer loss and society at large. Therefore, it is noteworthy to explore for both theoretical and practical interest the extent of techno-ethical behaviour an individual pursues and how the decision-making process is affected by various factors associated with technology usage. As unethical

conduct can be attributed mainly to individual behaviour (Jia, 2019), this study focuses on individual decision-makers and examines the role of their techno-ethical orientation in ensuring ethical decision-making (EDM).

Over the years, numerous studies (e.g. Ambrose *et al.*, 2008, 2021; Andreoli and Lefkowitz, 2009; Casali, 2010; Craft, 2013; Luca Casali and Perano, 2021; Malloy and Agarwal, 2010) have reported that an individual's ethical beliefs influence their approach to ethical or unethical behaviour, judgment and decision-making (Casali, 2010). It is proven that employees' ethical beliefs influence their decision-making process. Many studies in the existing literature propounded the role of individual characteristics (Deckers *et al.*, 2021; Haines and Leonard, 2007; Khalid *et al.*, 2012), moral identity (Arrami and QingXiang, 2021; Huhtala *et al.*, 2021; Resende and Porto, 2020; Wang *et al.*, 2021), cognitive moral development (Elm, 2019; Wittmer, 2019), moral affect (Arkin and Borenstein, 2019; Martinez and Jaeger, 2016; Zollo *et al.*, 2017), beliefs, attitude (Elbe and Brand, 2016; Kanjanapathy and Ramakrishnan, 2018; Zardeckaite-Matulaitiene *et al.*, 2014), etc., in influencing the decision-making process (Moore *et al.*, 2012; Moore, 2005). However, another stream of literature illustrates that individual-level factors cannot fully explain an individual's decision-making in an organisational context. The organisational setting that includes ethical climate, organisational culture, ethical leadership, organisational goal orientation, policies, regulations, members' acceptance, group identity and resources significantly drives the EDM. Similarly, technology is one of the resources for the organisation to conduct business. Therefore, technology does have a profound effect on EDM.

The guidelines for EDM have undergone a tremendous remodification in the context of the technological changes imposed due to the pandemic. The distinction between what constitutes ethical or unethical in the context of technology has become blurred (Verma and Garg, 2022). It requires revisiting the techno or digital theories to ensure an ethical workplace for the employees. In such a background, researchers advocate an adequate understanding of the nature of technology, referred to as technological frameworks, to make sense of how people derive ethical decisions. The technological frame (TF) of employees reveals how the group makes sense of the technology available in their organisation. These shared frames profoundly influence decision-making activity (Olesen, 2014). TFs may create predicaments in the process of EDM. Surprisingly, it shapes the way technology is constructed socially and shapes the way users are reconfiguring the technological utilisation.

Banking upon the social cognitive theory (SCT), this study examines the role of techno-ethical orientation and TFs in ensuring EDM in Indian supply chain companies. The present study attempts to structure the existing knowledge in two significant ways: (1) it examines the relationship between techno-ethical orientation and EDM and (2) it explores the moderating role of TFs in the relationship between techno-ethical orientation and EDM.

With the increasing concerns related to questionable business practices in organisations, this study investigates how an employee's technology-related behaviour and the meaning they associate with technology influence EDM processes. While there are shreds of evidence that suggest that organisations are ensuring an ethical code of conduct to prevent misuse of such practices by developing ethics awareness programmes, implementing jammers or blocking questionable websites during office hours, developing ethical leaders and ethical work climate, nonetheless, there is a lack of empirical evidence to substantiate the role of employee orientation and cognitive frames on EDM. By focusing on the relationship between techno-ethical orientation and TFs, this research is broadening the significant essence of employees or professionals working with technology to become ethically concerned with the role of technology in organisations and society at large (Batchelor *et al.*, 2012). Thus, the overarching premise is that TFs strengthen the interaction between a person's orientation towards ethical technology use and EDM in supply chain companies. Accordingly, the researchers of the current study synthesise the existing body of knowledge in resilience,

ethics and technology and advance the supply chain literature by answering the following research questions.

RQ1. How has technology facilitated the resilience in supply chain disruptions post COVID-19 pandemic?

RQ2. How do technology and ethics interact to shape EDM?

RQ3. How do individuals make sense of technology around them while making ethical decisions?

The paper is divided into three major sections. Section 2 reviews the existing literature and theoretical background related to technological and ethical interaction along with ethical principles and theories. Sections 3 and 4 provide a detailed description of the methodology, data collection approach and results. Finally, Section 5 discusses the findings and highlights the practical and theoretical implications.

2. Theoretical framework

2.1 Resilient supply chain

The supply chain industry remains prone to several risks and vulnerabilities. Unpredictable risk and disruption within the supply chain industry is a common pattern. The 2011 tsunami in Japan, the financial crisis of 2009, the computer chip factory fire in Japan, the 2003 SARS epidemic and others are notable examples of global events that exposed the supply chain receptivity to disruptions. Therefore, the topic of resilience always remains in discussion and is frequently addressed in every sector to contemplate the possible alternative for handling the disruptions. This time also, considering the damage caused in supply chain-related business operations, managers have turned their attention to better understanding and management of supply chain risk (Roh *et al.*, 2022). The industry has witnessed high-intensity pitfalls from production-linked damages to financial performance loss during the coronavirus outbreak. Recently, supply chain disruptions experienced across the globe varies from a drop in air travel by more than 95%, sudden panic buying of toilet papers, closure of warehouses and a sudden surge in hand sanitisers at an exponentially marked-up price. Nevertheless, the COVID-19 outbreak has revealed a particular type of supply chain disruption. Unlike earlier disruptions, this pandemic has impacted the supply chain more severely. It has affected all the nodes (supply chain members) and edges (ties) in a supply chain simultaneously (Chowdhury *et al.*, 2021). From significantly impacting food, healthcare, oil, airline, manufacturing and automotive industries to producers, suppliers, vendors, customers, government and countries; the repercussions of the coronavirus outbreak are multidimensional. In the pre-COVID-19 era, the researchers focused on a particular disruption scenario such as one involving supplier relations; however, during the coronavirus, the effect of a disrupted node propagated throughout the supply chain instantaneous. Moreover, the severe impact of the pandemic on India was completely unanticipated as China, the first country hit by the COVID-19 pandemic, serves as an essential exporter of various materials to India. The immediate consequence in China gradually crippled the supply chains in almost every sector (Kähkönen *et al.*, 2021). In the past, the companies tended to tolerate low-intensity regional epidemics effectively; however, for such extremely uncertain risk during this pandemic, the companies were unprepared for the distinct economic and financial challenges (Kähkönen *et al.*, 2021; Spieske and Birkel, 2021).

The managers' operational decisions usually focus on maximising profits, cost-efficient inventory, building a robust distribution network and return on investment. Instead, the industry requires strategic decisions to withstand the negative repercussions of the pandemic (Frederico *et al.*, 2021; Paul *et al.*, 2021). Survival of the firms can be achieved through strategic decisions such as building a culture of sustainability, creating plans for a smooth recovery

and leveraging the idea of maintaining a resilient supply chain. The idea of building a resilient supply chain can be viewed from multiple lenses. One school of thought asserted that the disposition of technology as a part of industry 4.0 and logistics 4.0 initiatives had paved the way forward for revolutionary transformation. Numerous articles have enumerated the importance of technologies for rethinking and managing the supply chain amid unexpected events (van Hoek, 2019; Rajesh, 2022; Wieland and Durach, 2021). Some authors have suggested that digitisation and deployment of specific modern technologies in the supply chain have significantly improved the decision-making process, thereby supporting the quality of response to pandemic-induced disruptions (Ivanov *et al.*, 2017; Lopes de Sousa Jabbour *et al.*, 2020). However, the scenarios mentioned above of recent disruptions have not only highlighted the need for building resilience but have questioned the ethical principle of the decision-makers. Researchers assert that in the recovery process, the individuals have neglected the ethical principles of their decisions. Undoubtedly, technology is seen as a competitive advantage for the industry and the only remedy to outpace the disruptions. However, the researchers of the present study have challenged the need to be "ethically resilient". Building upon the definition of resilience by several researchers (Hohenstein *et al.*, 2015; Katsaliaki *et al.*, 2021; Matyas and Pelling, 2015; Melnyk *et al.*, 2014; Wieland and Durach, 2021), the researchers of the current study proposed that being ethically resilient encompasses "a capability to adapt and proactively respond to any changes or disruptions with due regards to ethical principles leading to ethical decision-making".

2.2 Techno-ethics and techno-ethical orientation

Researchers have explored ethical perspectives concerning technology use under various terms such as netiquette, information ethics, digital ethics, techno-ethics, etc. These domains provide insights into the ethical dimensions incorporated within technological systems. Ethics of technology or techno-ethics is an interdisciplinary research area concerned with moral and ethical questions specific to the technology in society. The techno ethics approach involves integrating ICT practices and policies with ethical principles leading to ethically acceptable technology use. Techno-ethics derives from the moral and ethical belief system of an individual. Morality is a set of complex cognitive and emotional processes, every individual works according to their moral-cognitive schemas that shape their actions and behaviour.

Recently, researchers have reported an escalating unethical use of technology by students, workers, teachers and professionals across the globe (Kim *et al.*, 2022; Mât *et al.*, 2021; Otulugbu and Ogunobo, 2022; Smolinski *et al.*, 2022). Cyberbullying, infringement of privacy, plagiarism, deception, stealing information, cheating, fraud and defamation are a few challenges that erode technology's benefits. Researchers have stated that lack of techno-ethical training, generational differences, experience with technology, lack of resources and non-existence of technology-related ethical guidelines are primary reasons compelling individuals to engage in unethical practices (Camilleri, 2012; Verma and Garg, 2022). To appreciate the benefits of technology for society, it is crucial to analyse technology usage through the lens of ethical principles guiding the behaviour and actions of individuals.

2.3 Ethical decision-making

Decision-making is a critical process for an organisation's survival and success. Rationality and ethicality are two indispensable pillars of any organisational decision. Optimal outcomes emerge when leaders follow rational and EDM (Winkielman *et al.*, 2007; Zeni *et al.*, 2016). EDM processes are generally based on rationalist and non-rationalist approaches. The rationalist approach assumes that a decision-maker uses a rational, logical and deliberative cognitive process to conjugate moral standards (Schwartz, 2016). The most significant rationalist-based

theoretical models of EDM are Rest's (1979) four-component model for individual decision-making and Jones' (1991) issue-contingent model. Both EDM models emphasise the rational cognitive-moral process decision-makers use and their ethical orientation towards resolving ethical dilemmas. It implies that moral awareness and ethical belief systems determine the ethical decision. However, as the system becomes complex, decision-makers cannot rely only on their innate moral understanding. While arriving at an ethical decision, decision-makers consider individual, organisational and work environment factors (Herbert, 2010). They gather information or cues, process it and act according to their innate moral sense. The rationalist approach emphasises that various individual, organisational and moral intensity-related factors influence EDM (Lehnert *et al.*, 2015). The non-rationalist approach assumes that emotions (affective process) and intuitions (cognitive process) are central to moral judgment. One's "feeling state" (i.e. emotions) and "gut sense" (i.e. intuitive) are found to be more directly related to EDM. The non-rationalist approach places little emphasis on reasoning and more on people's behavioural adherence and emotional experience to their moral standards.

EDM at the workplace is gaining prime place because organisations, typically live in cyberspace which is dynamic, uncertain and vague. The perception of ethics in the digital world slightly varies from traditional definitions because, with technology and Internet, the recognition of right or wrong behaviour gets complicated (Bolhari *et al.*, 2017). When dealing with computers and the Internet, EDM is a fundamental factor that influences the successful implementation and utilisation of IT systems. Numerous researchers have argued the importance of addressing the ethical issues in technology and Internet usage because conventional theories (e.g. the theory of reasoned action) fail to satisfactorily model ethical decisions. Despite the best efforts by managers to ensure ethical practices in organisations, unethical activity by individuals is on the rise; hence, the importance of EDM will continue (De Cremer and Vandekerckhove, 2017; Zuber, 2015). Major moral theories, including utilitarianism, deontology and virtue ethics, have vigorously shaped EDM and provided sufficient evidence to measure what goes into making an ethical decision. Though numerous profiling instruments exist for EDM, most of them were developed and concentrated on moral cognitive aspects. For instance, defining issues test (DIT) by Rest (1979), corporate ethical values (CEV) by Hunt and Vitell (1986, 2006), and managerial judgement test (MJT) by Lind (1998) were designed around Kohlberg's theory of moral development; managerial value profile (MVP) by Sashkin *et al.* (1997) and multidimensional ethics scale (MES) by Reidenbach and Robin (1990) focus on categorisation of individual decision-making on ethical frameworks: utilitarianism, rights and justice. However, none of the instruments takes into account the self-virtues, or individual internal character traits (virtue ethics) – an essential aspect in influencing EDM (Casali, 2011; Crossan *et al.*, 2013; Luca Casali and Perano, 2021). This dimension of virtue ethics is central to developing an ethical environment in any business or organisation. Virtuous traits such as equity, confidentiality, trust, compassion, loyalty, fairness and openness shape a person's and an organisation's vision thereby, enabling unity of character and integrity in a person (Audi, 2012; Shanahan and Hyman, 2003). Unlike other ethical principles, virtue ethics are intrinsic to business ethics that apply both deontological and teleological criteria when making ethical judgements. Businesses often apply and follow several universal virtues of autonomy, duty, equity, integrity, loyalty, legality, entrepreneurialism, honesty, honour, reliability, responsibility, etc. Similarly, scholars have argued that technology (an important resource for the organisation) must exhibit the same virtues to ensure techno-ethical decisions. A virtuous-oriented person aims at the right target and deepens his character strengths along with the right means through self-reflection (reflection-in-action). In this connection, an individual's reflection upon one's own or others' ethical or unethical conduct related to technology usage can foster techno-EDM. Hence, the optimum utilisation and successful implementation of IT systems in the organisation can be achieved. In turn, these practices will give rise to ethical organisation.

2.4 Techno-ethical orientation and ethical decision-making

The proposed relationship between techno-ethical orientation and EDM is theoretically premised on SCT (Bandura, 1988, 2011). The theory states that individuals' actions and motivation are regulated by forethought. Self-efficacy, outcome expectations, goals and socio-structural factors are the core constructs of the SCT that influence one's behaviour. SCT also provides answers to moral and ethical orientation and decision-making (Claybourn, 2011; Galperin *et al.*, 2011). This theory adopts a cognitive interactionist perspective on moral and ethical phenomena. Generally, SCT states that social and cognitive aspects influence behaviour.

The social aspect of SCT focuses on an individual's actions being influenced by the environment and extraneous factors and conditions. It refers to an individual's capacity to evaluate one's behaviour with environmental moral standards and set the future course of action by comparing past behaviour outcomes and moral standards. The cognitive aspect focuses on personal attributes and thought processes influencing individuals' behaviour. One's ethical orientation is regulated by self-assessment of their standard of conduct within environmental circumstances. As long as the gap between one's ethical behaviour and environmental moral standards is less, the action will be deemed righteous. And, if the moral standards of the environment are far from one's cognition and thought process, those actions will be considered unethical.

Based on these considerations, SCT helps in analysing factors that influence EDM. It considers both society's moral standards and one's ethical principles and how they interact while arriving at an ethical decision. In particular, employees' behaviour towards technology is contingent upon rules and standards prevailing in the organisation, such as code of ethics and netiquette, and their expectations based on how the technology will serve individual needs. Thus, one's techno-ethical orientation is poised to influence their decision-making prospects. Therefore, it is appropriate to state that people's attitude/orientation towards technology and its consequences, whether ethical or unethical, determines their moral judgment and decision-making. In support of this, it is proposed that an individual's techno-ethical orientation influences the EDM in the organisation, and accordingly, the following hypothesis is formulated:

H1. Techno-ethical orientation is positively associated with EDM

2.5 Technological frames as a moderator

Individuals use specific frames while using technology to make sense of it. These are called TFs (Orlikowski and Gash, 1994). TF refers to an individual's assumptions, knowledge and expectations regarding the ethical impact of introducing and using certain technology at the individual, organisational and cultural levels.

These frames are the foundation through which "we filter and then interpret the actions of others and our environment to make sense of our world" (Olesen *et al.*, 2013). In an organisational setup, as socialisation or interaction occurs among members of the organisation, they develop common and shared frames in each department (Van Hulst and Yanow, 2016). Orlikowski and Gash posited that TFs are assumed to be shared by individuals. And these shared frames guide the members' interactions around technology and their understanding and use of technology. Moreover, as the concept of TF is derived from SCT, participants' frames and sense-making play a fundamental role in setting a decision criterion for adopting and using new technology.

From a socio-cognitive perspective, TFs and sense-making (i.e. an individual's ability to attribute meaning to new, ambiguous or equivocal situations or actions) are fundamentally social processes that apply to individuals and groups. They reflect the meaning and interpretations of people's actions and how they organise their decisions. Therefore, TFs are

triggered when participants encounter uncertainty or ambiguity in their core meaning systems or identity while using technology. In line with the idea that TF is a part of social discourse and conditioned by cognitive structures, the notion of EDM is governed by the understanding of members' cognitive frames and the meaning they associate (Esmaili *et al.*, 2021; Al Halbusi *et al.*, 2017; Lin *et al.*, 2018; Yun *et al.*, 2020).

Furthermore, individuals' ethical or unethical use of technology can be exacerbated by a changing TF of reference. With the passage of time and advancement in technical systems, the frames changes and so does the usage pattern by individuals (Tyre and Orlikowski, 1994). Individuals modify their frames or knowledge structure to make way for the effective utilisation of technologies and devices. In addition, effective utilisation of technology in an organisation depends on the congruence or incongruence of individual and organisational member (group) TF. If alignment (congruence) exists between structure and content, it contributes to the effective utilisation of technology by organisational members. And, if incongruence is present in the frames of key organisational stakeholders' groups, problems such as resistance and unethical use of technologies may result (Orlikowski and Gash, 1994). It will be appropriate to state that understanding and interpretation of TFs would be beneficial to ensure EDM. In this essence, researchers propose that EDM in organisations is governed by members' cognitive structures or TFs, i.e. their assumptions, beliefs and expectations concerning technology usage.

The theory of TFs identified three domains of TFs relevant to the adoption of technology in a group: (1) *the nature of technology*-people's image and understanding of technology's capabilities and functionality; (2) *technology-in-use*- people's understanding of how the technology will be used and consequences associated with such use; and (3) *technology strategy*-people's view of motivation or vision behind organisation's adoption of technology. These three categories determine employees' perception of technology and are likely to enhance their interaction, influencing the organisation's decision-making process (Criado and O.de Zarate-Alcarazo, 2022). Building upon these categories, previous studies have directly assessed TF as a multi-item construct. Mishra and Agarwal (2010) differentiated between benefit, threat and adjustment frames. DeVellis (2012) tried to measure TF on a single-item measure. Spieth *et al.* (2021) recently attempted to measure TF through a multi-item scale. Their scale explored the variety of individual reactions to technologies implemented by firms on five dimensions: personal attitude, application value, supervisor influence, organisational influence and industry influence. All these dimensions influence the decision-making process. The present study is premised on this conceptualisation of TFs.

An employee's attitude towards an option or stimulus serves as a primary basis for evaluating and assessing his actions. For instance, the resultant outcome would also be positive for an individual with a positive outlook on technology. A person's ethical orientation and personal attitude influence his ethical decisions. As a consequence, the hypothesis in this support is as follows:

H2. Personal attitude towards technology moderates the relationship between techno-ethical orientation and EDM.

As premised in SCT, the socio-structural factors facilitate or impede the actions or behaviour. Researchers propose that if the technology can facilitate coordination of work, provide flexibility and reduces the possibility of making mistakes, the decision-making process concerning technology usage would be positive and ethical. Application value strengthens the interaction between techno-ethical orientation and EDM. The hypothesis for the argument is as follows:

H3. Technology's application value facilitates the relationship between techno-ethical orientation and EDM.

EDM is regulated by environmental factors such as organisational policies, organisation climate, supervisor actions and peer/group ethical schema. An organisation's management support facilitates ethical behaviour by employees. The organisation's and its members' support facilitate effective employee engagement with ICT (O'Driscoll *et al.*, 2010). The organisation's inclination towards implementing technologies at the firm and its effort in providing assistance and training to participants enable better and more EDM at the organisation. Researchers have supported that general organisational support is associated with positive employee outcomes such as greater job satisfaction, better employees' health and well-being, enhanced productivity and better decision-making. According to social exchange theory, employees are motivated and produce high-quality output when they feel that the organisation supports them (Park *et al.*, 2018). For instance, if employees in supply chain industries are provided with sufficient and appropriate training to work with technology, they are more likely to reciprocate with high-quality value addition to the firm in terms of profitability, job engagement, employee retention, etc. Therefore, this study assumes that organisational factors facilitate the association between techno-ethical orientation and EDM. The proposed hypothesis is as follows:

H4. Organisational factors facilitate the relationship between techno-ethical orientation and EDM.

Also, previous researchers have reported that people tend to engage in poor ethical practices if their superiors or peers are not ethical (Odole, 2018; Ünal *et al.*, 2012). Supervisors' willingness and behaviour regarding technology use determine the interaction between an individual's techno-ethical orientation and EDM. Support from supervisors has been suggested as a powerful tool to enhance supply chain employees' attitudes and behaviours, including technology usage and decision-making. Especially, supervisors are expected to play a vital role in encouraging the employees to work with a new approach, support them during the transition period from traditional techniques to modern and tech-enabled techniques and guide them in applying the competencies that they have learned in the workplace while ensuring job security for the employees (Park *et al.*, 2018). Recent studies have found that managers' support mediates the relationship between organisational support and innovative behaviour (Crowley and Bourke, 2018; Gözükara and Çolakoğlu, 2015; Huo *et al.*, 2020; Zhang *et al.*, 2008). It states that employees are more likely to get influenced by the supportive approach of the managers. Likewise, in supply chain industries, studies have emphasised the facilitating role of senior managers or middle managers in the implementation of innovation and quality development initiatives (Engle *et al.*, 2017; Sönmez and Yıldırım, 2019). Accordingly, the following hypothesis framed is as follows:

H5. Supervisor's support facilitates the relationship between techno-ethical orientation and EDM.

Similarly, with the automation and development in technology and ICT devices, the working of the industry ecosystem smoothens and advances. Moreover, researchers have reported differences in technological usage and adoption across industries (Kutnjak *et al.*, 2019; Oh and Shim, 2020; Shrestha, 2021). Industries using IT systems have reported growth in operating profit margins and higher labor productivity (Bessen, 2017, 2020; Franceli and Zilber Turri, 2021). The extant literature has shown that deploying IT and using IT knowledge in managing the supply chain is one of the prime success factors (Basheer *et al.*, 2016; Nasiri *et al.*, 2020). Higher the IT capability of a supply chain industry, the greater the resilience and performance of the industry during uncertain times. Consequently, the industry significantly influences the relationship between techno-ethical orientation and EDM. The proposed hypothesis is as follows:

H6. Industry factors facilitate the relationship between techno-ethical orientation and EDM.

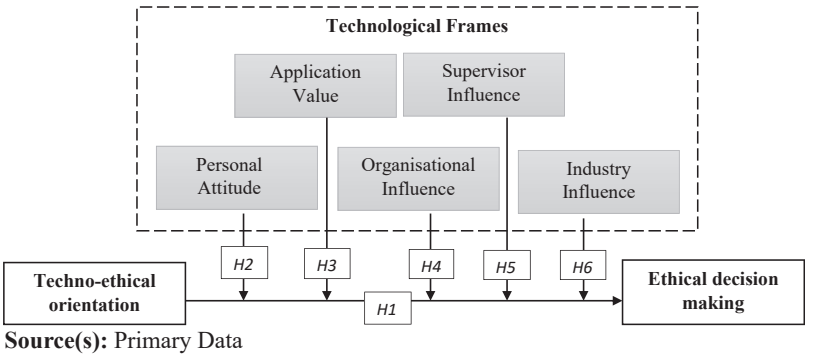
In a nutshell, the present study posits that techno-ethical orientation and TFs (i.e. personal attitude, application value, organisation influence, supervisor influence and industry influence) play an important role in ensuring EDM. And the interaction between techno-ethical orientation and EDM is moderated by TFs. Figure 1 presents the research model and hypotheses.

3. Research methodology

The present study employs an exploratory research design. The exploratory research design is categorised by flexibility and versatility concerning the methods because it does not entail the use of formal research protocols and procedures. The focus of exploratory research is to discover new insights and provide more information on the subject to be investigated (Malhotra, 2016). For this reason, exploratory research is meaningful in any situation where the researcher does not have enough understanding to proceed with the research project (Provdanov and De Freitas, 2013).

3.1 Sample and data collection

The data were collected through an online survey using cross-sectional and purposive sampling methods. Inclusion criteria for the study required that participants be engaged in the supply chain industry and have substantial experience working with technology for business operations. A total of 388 responses were received through emails and social media platforms. Forty-two responses were discarded for being incomplete and invalid. The sample included 346 employees of supply chain companies in India. The survey questionnaire was shared and posted on various online platforms such as Google Groups, Yahoo Groups, LinkedIn and some companies' internal discussion groups and forums for maximum reach and responses. Table 1 presents the demographic information of the respondents. Among these respondents, 45.3% were in the age group of 25–35 years, followed by 21.0% below the age of 25 years and 16.7% in the age group of 45–55 years. Also, 69% of the participants had more than five years of technology experience in their work life. Gender-wise, 66% were males and 34% were females. The participants' marital status and working experience are presented in Table 1. In total, 44% of the participants had work experience between 5 and 10 years in supply chain companies, followed by 32% of participants with more than 10 years of work experience. Also, almost half of the participants were married, i.e. 55%, and 45% were unmarried.



Source(s): Primary Data

Figure 1.
Research model

Variable		Frequency	Percentage (%)	Cumulative (%)
Gender	Male	227	66	66
	Female	119	34	100
	Total	346		
Age	Below 25 years	73	21.0	21.0
	25–35 years	157	45.3	66.3
	35–45 years	44	13.0	79.3
	45–55 years	58	16.7	96
	Above 55 years	14	4.0	100
	Total	346		
Marital status	Married	190	55	55
	Unmarried	156	45	100
	Total	346		
Work experience	Less than 5 years	83	24	24
	5–10 years	152	44	68
	More than 10 years	111	32	100
	Total	346		
Experience with technology	Less than 3 years	22	6	6
	3–5 years	85	25	31
	More than 5 years	239	69	100
	Total	346		

Table 1. Demographic characteristics of respondents ($N = 346$)

Source(s): Primary Data

3.2 Common method bias

As the data were collected from a single survey at one time, the issue of common method bias may be a concern in this study. Podsakoff *et al.* (2003) suggested procedural and statistical control remedies avoid the common method biases. Firstly, relying upon a single format for either the scale or scale anchors is one of the major sources of common method bias. In order to minimise the impact of common method variance, the constructs were measured on a different point scale. In addition, psychological separation of measurement was created by using a cover story to make it appear that the measurement of the independent variable is not connected with or related to the measurement of the outcome variable. This technique reduces the respondent's perceived relevance of the previously recalled information or their prior responses to answer subsequent questions. It reduces the consistency in their responses (Podsakoff *et al.*, 2003). Second, the marker variable technique (Lindell and Whitney, 2001) was used as a statistical approach to test for common method bias. Essentially, the marker variable technique uses a variable theoretically unrelated to any substantive variables being studied to calculate the influence of common method variance and adjust the correlations among all constructs (Jordan and Troth, 2020; Schwarz *et al.*, 2017). Researchers incorporated a six-item marker variable-Gratitude (GQ-6) to assess individual differences in the proneness to experience gratitude in daily life (McCullough *et al.*, 2002). Results showed the lowest correlation between the marker variable and the remaining variables, and the obtained adjusted correlations are significant. Thus, it is evident that common method bias was not a pervasive issue.

3.3 Instruments

3.3.1 Techno-ethical orientation. A 14-item scale developed and validated by Verma and Garg (2022) was used to measure the techno-ethical orientation of the employees across four dimensions, namely, infringement of the right to privacy, self-enrichment, defamation and loafing during work hours was used. The responses were collected on a five-point scale

ranging from 1 (highly unethical) to 5 (highly ethical). Sample items included "Not giving due credits to someone for providing project-related material," "Copying software from the organisation for personal use," and "Discussing negative aspects of the colleagues or organisation with others on the Internet" (see Appendix).

3.3.2 Ethical decision-making. A 24-item scale developed by Casali (2011) was used to capture the profile of EDM styles on a five-point Likert scale ranging from 1 (extremely important) to 5 (not important at all). Some sample items included "respecting dignity of those affected," "giving the opportunity to all affected parties or their representatives to have input into the decision-making process" (see Appendix).

3.3.3 Technological frames. A 20-item scale developed by Spieth *et al.* (2021) capturing the dimensions of TF construct in the digital age on a seven-point scale ranging from 1 (strongly disagree) to 7 (strongly agree) was used. The word "digital technologies" from the original scale was replaced with the word "technologies" to suit the purpose of this study. Sample items included "Technologies reduce the possibility of making mistakes in my work," "Technologies are an important part of my life," and "Our customers demand the use of technologies" (see Appendix).

3.4 Reliability and validity

The reliability of techno-ethical orientation, EDM and different dimensions of the technological framework was measured with the help of Cronbach's alpha and composite reliability (CR) values. The Cronbach's alpha (α) and CR estimates (Table 2) were greater than 0.70, which established the reliability of the study variables (Hair *et al.*, 2010; Henseler *et al.*, 2009). In addition, the average variance extracted (AVE) values of all constructs were higher than the recommended threshold of 0.50, which concluded the convergent validity of the scale (Fornell and Larcker, 1981). The discriminant validity was tested for all the five components of the moderator "Technological Frames" using Fornell and Larcker's (1981) criteria. The square root of AVE values (highlighted in bold and placed diagonally in Table 3) is higher than its correlation with other constructs. It confirms the discriminant validity of the scale.

3.5 Data analysis

Structural equation modelling (SEM) was used to test the theoretical model and hypothesis proposed in this study. SEM is considered a large sample method (minimum sample size of 200) that offers flexibility and generality to evaluate any theoretical model (Hayes, 2018; Hayes *et al.*, 2017). The use of SEM for the present study is considered appropriate given the sample size and the measurement model proposed for research. IBM SPSS AMOS (Analysis of Moment Structure) software accurately reflects complex relationships in attitudinal and behavioural models through various multivariate analysis methods (IBM Corp., 2016).

Construct	Mean	Std. deviation	α	CR	AVE
Techno-ethical orientation	2.93	1.82	0.917	0.955	0.687
Ethical decision-making	3.38	0.55	0.899	0.921	0.655
Personal attitude	5.83	0.99	0.885	0.842	0.572
Application value	5.71	0.91	0.894	0.858	0.595
Organisational influence	4.25	1.48	0.929	0.894	0.631
Supervisor influence	5.10	1.13	0.869	0.775	0.655
Industry influence	5.16	1.27	0.939	0.899	0.607

Note(s): α – Cronbach's alpha, CR – Composite Reliability, AVE – Average Variance Extracted
Source(s): Primary Data

Table 2.
Descriptive statistics

Since the present study requires testing and confirming existing theoretical models, covariance-based-SEM (CB-SEM) is considered an appropriate method than PLS-SEM. Moreover, Amos is more stringent than Smart PLS (Hair *et al.*, 2017). One of the advantages of CB-SEM over PLS-SEM is the tendency of the PLS to cause overestimation of the measurement parameters and the underestimation of the structural parameters (Al Issa and Abdelsalam, 2021). For the current study, the AMOS standard estimation model, i.e. maximum likelihood estimation (MLE), was employed. The basic assumptions for running MLE, such as large sample size, multivariate normal distribution and validity of the hypothesised model, were fulfilled. For normality check, Tabachnick and Fidell (2007) recommended the normal range for skewness-kurtosis value. When the skewness-kurtosis statistic is divided by its standard error and the value is greater than $z = \pm 3.29$ ($p < 0.001$, two-tailed test), it raises normality concerns. However, all the items in the dataset were found to be normally distributed (i.e. less than $z = \pm 3.29$). More specifically, the skewness and kurtosis value in each case was less than ± 1, indicating normality in the dataset of this study. Hair *et al.* (2010) recommended the existence of multicollinearity if the inter-correlation amongst the constructs was more than 0.90. The inter-correlation among the constructs for this study ranged from 0.289 to 0.753; thereby, no evident issue of multicollinearity was found.

4. Results

Table 2 illustrates descriptive statistics, reliability and validity estimates of study variables. The relationship between techno-ethical orientation and EDM was analysed using correlation and regression. Tables 4 and 5 demonstrate correlation and regression values between techno-ethical orientation and EDM. The correlation coefficient between independent and dependent variables was 0.513 (Table 4). Also, as indicated in Table 5, the regression coefficient ($\beta = 0.213$) was reported significant at 0.05, which provides empirical support for

Table 3. Measure of discriminant validity

Construct	1	2	3	4	5
Personal attitude	*0.756*				
Application value	0.743	*0.771*			
Organisational influence	0.289	0.432	*0.794*		
Supervisor influence	0.442	0.408	0.669	*0.809*	
Industry influence	0.515	0.493	0.555	0.678	*0.779*

Note(s): Square root of AVE values in diagonal and highlighted in italic
Source(s): Primary Data

Table 4. Correlation among variables

Variables	1	2	3	4	5	6	7
TEO	1						
PA	0.423	1					
AV	0.381	0.753*	1				
OI	0.327	0.289	0.432*	1			
SI	0.431	0.442*	0.408*	0.669*	1		
II	0.362	0.515*	0.493*	0.555*	0.678*	1	
EDM	0.513*	0.303	0.464	0.387*	0.463	0.305	1

Note(s): *Sig at 0.05, TEO – Techno-ethical orientation, PA – Personal attitude, AV – Application value, OI – Organisation influence, SI – Supervisor influence, II – Industry influence, EDM – Ethical decision-making
Source(s): Primary Data

the acceptance of the first hypothesis. *R*-square values showed a 45.2% variation in EDM is explained by techno-ethical orientation. It suggests that the techno-ethical orientation is positively associated with EDM. The moderation effects of the five dimensions of TFs were analysed through SEM in AMOS v24.0. Accordingly, a moderation model was developed. The proposed measurement and structural model factors showed good model fit indices with accepted threshold limits recommended by Hair *et al.* (2010) and Sahoo (2019). The results of the confirmatory factor analysis are presented in Table 6.

4.1 Moderation analysis

Tables 7–11 explain the interaction effect of moderators and the independent variable (techno-ethical orientation) on the dependent variable (EDM). The results indicate a positive relationship between techno-ethical orientation and EDM. And this relationship is strengthened in the presence of personal attitude (Table 7). Also, the moderating impact of personal attitude on the relationship between techno-ethical orientation and EDM style is significant ($b = 0.154$, $t = 2.42$, $p = 0.015$). Figure 2 stated that the line for high PA is much steeper, indicating that with a high personal attitude (positive attitude towards technologies), the impact of techno-ethical orientation on EDM style is much stronger than low PA. As long

IV	DV	Unstandardized co-efficient	Standard error	Standardized co-efficient	*R*-square	*p*-value
TE	EDM	0.342	0.236	0.213	0.452	0.015*

Note(s): *Sig at 0.05, TE – Techno-ethical orientation, EDM – Ethical decision-making, IV – Independent variable, DV – Dependent variable
Source(s): Primary Data

Table 5. Summary of Regression results

Fitness indices	Measurement model	Structural model	Cut-off criteria
CMIN/df	4.25	3.33	2–5
CFI	0.901	0.922	≥0.90
RMSEA	0.05	0.04	<0.06
p-value	0.42	0.06	>0.05
SRMR	0.057	0.070	<0.08
TLI	0.952	0.956	≥0.95
GFI	0.955	0.958	≥0.95

Note(s): CMIN/df – Chi-square/degree of freedom, CFI – Comparative fit index, RMSEA – Root mean square error of approximation, SRMR – Standardized root mean squared residuals, TLI – Tucker–Lewis index, GFI – Goodness-of-fit index
Source(s): Primary Data

Table 6. Model fitness indices

Relationship	Beta	C.R	*p*-value
Techno-ethics → Ethical decision-making	0.564	2.34	0.019*
Personal attitude → Ethical decision-making	0.041	0.08	0.935
Techno-ethics × Personal attitude → Ethical decision-making	0.154	2.42	0.015*

Note(s): *Sig at 0.05
Source(s): Primary Data

Table 7. Moderation analysis summary for personal attitude as moderator

as the attitude towards technologies is positive, the orientation and individual's decision-making style are more ethical. Thus, hypothesis H2 is supported.

Similarly, the moderating impact of application value on the relationship between techno-ethical orientation and EDM is shown in Table 8. The results indicate a positive relationship between techno-ethical orientation and EDM. And this relationship is strengthened in the presence of application value (Table 8). Also, the moderating impact of application value on the relationship between techno-ethical orientation and EDM style is significant ($b = 0.183$, $t = 2.40, p = 0.016$). It can be stated that the line for high AV is steeper, indicating that at high application value (technology use is valuable), the impact of techno-ethical orientation on EDM style is much more substantial as compared to low AV (Figure 3). As long as the technology provides valuable benefits in work, the orientation and decision-making style are more inclined toward ethical technology-related behaviour. Surprisingly, the

Table 8. Moderation analysis summary for application value as moderator

Relationship	Beta	C.R	*p*-value
Techno-ethics → Ethical decision-making	0.617	2.44	0.015*
Application value → Ethical decision-making	0.228	0.430	0.667
Techno-ethics × Application value → Ethical decision-making	0.183	2.40	0.016*

Note(s): *Sig at 0.05
Source(s): Primary Data

Table 9. Moderation analysis summary for organisational influence as moderator

Relationship	Beta	C.R	*p*-value
Techno-ethics → Ethical decision-making	0.425	1.735	0.083
Organisation influence → Ethical decision-making	0.258	0.590	0.112
Techno-ethics × Organisation influence → Ethical decision-making	0.501	1.853	0.046*

Note(s): *Sig at 0.05
Source(s): Primary Data

Table 10. Moderation analysis summary for supervisor influence as moderator

Relationship	Beta	C.R	*p*-value
Techno-ethics → Ethical decision-making	0.352	0.506	0.013*
Supervisor influence → Ethical decision-making	0.179	0.398	0.691
Techno-ethics × Supervisor influence → Ethical decision-making	0.540	0.916	0.030*

Note(s): *Sig at 0.05
Source(s): Primary Data

Table 11. Moderation analysis summary for industry influence as moderator

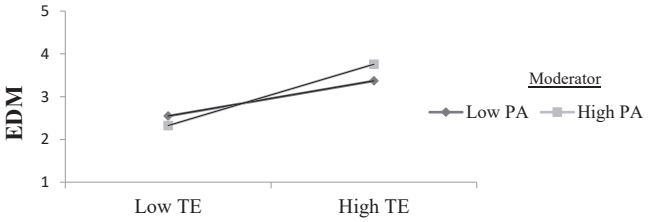
Relationship	Beta	C.R	*p*-value
Techno-ethics → Ethical decision-making	0.355	−1.515	0.130
Industry influence → Ethical decision-making	0.008	0.020	0.934
Techno-ethics × Industry influence → Ethical decision-making	0.401	−0.568	0.050*

Note(s): *Sig at 0.05
Source(s): Primary Data

decision-making style and orientation were still ethically inclined at low AV. Thus, hypothesis H3 is supported.

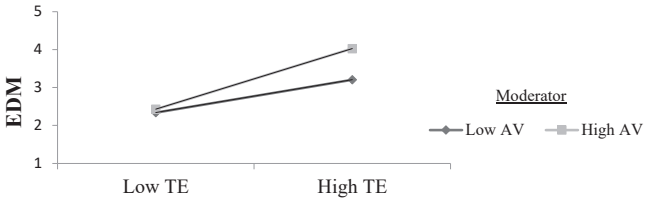
Apart from personal and technology-linked factors, the role of organisation support, supervisor support and industry factors during technology usage were assessed (Tables 9–11). At the organisational level, colleagues' and organisations' support while using technology reinforces the positive relationship between techno-ethical orientation and EDM (Table 9). In the presence of organisational factors such as help from colleagues in using technologies for the job, the relationship between DV and IV strengthens and becomes significant ($b = 0.501$, $t = 1.853$, p-value $= 0.046$). Therefore, managers' decision-making style depends on organisational factors (Figure 4). Thereby, hypothesis H4 is supported.

The supervisor's support and willingness to integrate the technology with the firm has a significant and positive impact on influencing the relationship between a person's techno-ethical orientation and EDM ($b = 0.540, t = 0.916, p$ value $= 0.030$) (Table 10). If the supervisor requests the use of technology and regularly supports its implementation, the orientation and decisions related to technology will be more ethical. Higher the supervisor's support, highly ethical the orientation and decision-making style will be (Figure 5). Thereby, hypothesis H5 is supported.



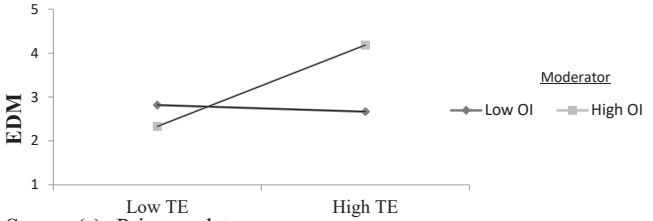
Source(s): Primary data

Figure 2.
Moderation effect of personal attitude (PA) on techno-ethical orientation (TE) and ethical decision-making (EDM)



Source(s): Primary data

Figure 3.
Moderation effect of application value (AV) on techno-ethical orientation (TE) and ethical decision-making (EDM)



Source(s): Primary data

Figure 4.
Moderation effect of organisational influence (OI) on techno-ethical orientation (TE) and ethical decision-making (EDM)

The data at the industry level were also analysed, and the results for the same are presented in Table 11. Industry factors such as demand from competitors to use the technologies have a significant and major influence on a person's techno-ethical orientation and EDM ($b = 0.401$, $t = -0.568$, $p = 0.050$). Industry factors strengthen the positive relationship between IV and DV, and this impact is noticeable (Figure 6). Hypothesis H6 is supported (positive and significant relationship).

5. Discussion

Technology is the backbone for business organisations and global markets to systematically integrate their operations and connect with customers even in a complex and dynamic environment. With the emergence of the IoTs, big data and cloud-based systems, industries have revolutionised their operations, especially in the manufacturing and supply chain sectors. Due to technological innovations and state-of-art infrastructure, the supply chain industry saw impressive growth in its operation and interaction (Björkdahl, 2020; Liu and Chiu, 2021). However, the COVID-19 outbreak has significantly affected domestic and global supply chain operations (Paul *et al.*, 2021). The disruption in transportation and distribution systems, supply of products and reduced finances, questions about the survivability and sustainability of the supply chain network. The practitioners and decision-makers were not ready to manage the pandemic's severe impact, and the supply chain's resilience was compromised (Paul *et al.*, 2021). Despite these challenges, this industry recovered effectively with the help of high-tech and specialised technologies even post-COVID-19 era. However, employees and staff struggle to adjust to new processes and tools due to a lack of training, practice and knowledge about software, the Internet and ICT devices. Not everyone finds himself comfortable with new technologies. It leads to growing incidents of wrongful and unethical practices with technology. Given the potential and substantiate contribution of the supply chain sector to a country's economic and financial growth, assessing an individual's techno-ethical orientation and its influence on EDM is imperative. To analyse this in detail,

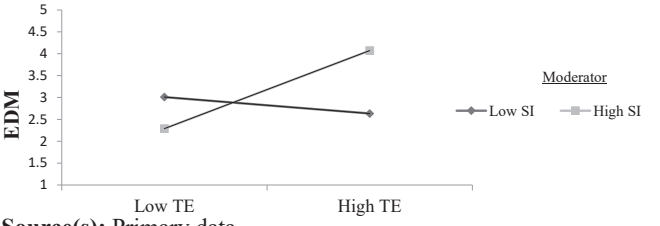


Figure 5.
Moderation effect of supervisor influence (SI) on techno-ethical orientation (TE) and ethical decision-making (EDM)

Source(s): Primary data

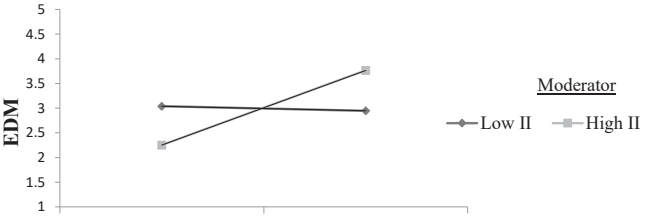


Figure 6.
Moderation effect of industry influence (II) on techno-ethical orientation (TE) and ethical decision-making (EDM)

Source(s): Primary data

the present study focuses on evaluating the effect of an individual's techno-ethical orientation on EDM in the presence of TFs.

The primary objective of this study was to examine the relationship between the techno-ethical orientation of employees and the EDM process in the supply chain industry in the presence of various factors influencing the relationship. The research study examines the moderating role of TFs (personal attitude, application value, organisational factors, supervisor influence and industry influence) on the proposed association between techno-ethical orientation and EDM. The proposed research model was established based on the extant literature. Six hypotheses were consequently confirmed. Drawing from the empirical results using SEM Amos, all six hypotheses were statistically supported and found to be significant. Techno-ethical orientation has a positive and significant influence on EDM.

Similarly, all variables of TFs have a positive and significant moderating effect on the relationship between techno-ethical orientation and EDM. The study results signify that having a positive attitude towards technology, valuable services offered by technology, support from organisations, supervisors and increasing demand for technology usage at the industry level enables a decision-maker to utilise technology ethically and improves decision-making. The investment in digital tools, technology and EDM will help in maintaining an ethical supply chain in the era of logistics 4.0.

5.1 Theoretical implications

SCT has underpinned several studies in IT, ethics, psychology and science. Notably, the analysis of this study is considered one-of-a-kind to evaluate the application of SCT in supply chain management. Given the pervasive Internet use and technology in supply chain management, this study's analysis revealed the need for a digital and ethical supply chain. The study contributed to creating new insights into four areas of interest to the research community: (1) supply chain resilience; (2) techno-ethics; (3) ethical decision-making; and (4) TFs. This study confirms the finding of the majority of the relationships between the constructs as suggested in the past literature (e.g. the role of IT in building a resilient supply chain and the interaction between technology and ethics that shapes EDM). Moreover, the framework of TFs has not been studied in the ethical domain. The study results supported that the frames they used influence persons' ethical orientation towards technology. In addition, the findings also propose the need to be "ethically resilient" owing to the unforeseen risk caused by the COVID-19 pandemic. Therefore, this study provides a new dimension to consider the role of ethics when a decision-maker makes sense of the technology around him. As survival of an organisation is impossible without technology, TFs shape technology trajectories over the entire technology lifecycle (Kaplan and Tripsas, 2008).

5.2 Practical implications

The study results are significant for management, practitioners and industry experts. COVID-19 has disrupted both the upstream and downstream supply chains and has exposed the ill-preparedness of the supply chains and logistics sector to cope with such adverse situations. In this context, the present study's results will enable managers to develop strategies for mitigating employee risk and unethical practices. Digital supply chain and logistics 4.0 are inevitable. Investing in the right technology is quintessential for increasing the supply chain network's robustness, resilience and sustainability (Haddud and Khare, 2020). The question is how to ensure that technology usage in the industry is ethical. Knowing one's ethical orientation towards technology allows managers to instigate EDM in the organisation, focusing on sustainability, resilience and ethical supply chain. Recovering quickly to one's original state from an unknown disruption or turbulent change requires careful deliberation in adopting reactive or proactive strategies. Adopting transparent

technologies that can monitor and streamline the supply chain processes ethically at the expense of increasing unethical technology-related practices is insufficient to justify. Therefore, decision-making is a pivotal process for the long-term establishment of an organisation and the proper functioning of business activities. Previous researchers had established the role of EDM on a firm's performance, employee well-being, commitment, job satisfaction, etc. But the results of this study will enable managers to consider the role of social and organisational factors in influencing ethical choices. For successful implementation and utilisation of technology, managers have to ensure that those technologies provide positive benefits to the users; employees are comfortable using technology and the management supports the use of technology. All factors taken together will enhance the experience of technology users, and as a consequence, this will lead to techno-ethical judgment, culture and decision-making. In addition, the results of the study will be beneficial for businesses and policymakers in the supply chain in the following ways: (1) Supply chain companies in India should set up a continuous monitoring mechanism to track unethical practices in the firm that can severely tarnish its image as a leading contributor to the global supply chain. Such surveillance and monitoring tracker can help identify the wrongdoers before they worsen the firm's goodwill, and steps could be taken to implement an ethical code of conduct concerning technology usage; (2) A strategy and policy to regulate and govern the implementation and utilisation of modern technology such as blockchain, cloud computing, big data, IoTs, business intelligence, digital supply chain, smart robots, etc., should be expedited on an urgent basis to combat a further surge in unethical practices; (3) Since employees fear a loss of their job due to the introduction of technology and their incapability to adopt working with new technology, appropriate training and development initiatives should be encouraged for employees at all levels. Technology uncertainty is one of the leading reasons for rising unethical conduct. Nevertheless, organisations should plan training to get rid of uncertainty. Teaching optimum utilisation of technology can foster creativity, innovation and value addition to the firm through an employee ethical code of conduct; (4) Auditing the suppliers from time to time will allow the supply chain industry to ensure that they operate ethically. This way, suppliers will comply with company goals, and in turn, this will improve product quality and enrich the supplier-vendor relationship; (5) Rewarding suppliers and employees – providing incentives to those who perform ethically can be a useful strategy for developing a culture of ethics.

6. Limitations and scope

The findings of the study are subject to several limitations. Firstly, as this study is based on a cross-sectional survey approach, a longitudinal research design could have produced more insights into how the frames are shaped and techno-ethical orientation evolves with newer technologies. A comparative study before and during the introduction of the digital supply chain would provide a better idea of how an employee behaves and responds to technological innovations. Moreover, we recognised the use of a mixed-method approach for future studies that would help researchers and practitioners to explore the viewpoints of employees regarding working and learning from the latest technologies.

7. Conclusion

The supply chain industry is highly vulnerable yet essential to keep the economy going. The supply chain maintains a balance in the trade and economy of a country. When the industry is hit with disruptions such as the recent global health crisis, i.e. coronavirus, it needs to be recovered effectively to repair the damages that occurred. During such time, the technology

has proven to be a privilege for the industry to become resilient and recover productively. With technological advancement, the supply chain industry received immense benefits and responded proactively to the pandemic-induced disruptions. However, with the growing development of technologies paving the way for revolutionary transformation, the decision-making related to technology use is indistinct. Ethical principles are often overlooked while working with technology. The researchers here proposed the need to be "ethically resilient" in the course of mitigating the risk and disruptions. And technology, being a part of this journey, needs to be evaluated in terms of ethical principles for EDM. The present research addresses this gap and identifies the novel interrelationship between an individual's techno-ethical orientation and ethical decision-making in the presence of TFs as a moderator. Consequently, the findings suggested that one's techno-ethical orientation translates to EDM when it can be determined how individuals make sense of technology around them (i.e. TFs). The present study is one of the pioneer studies exploring the interaction of technology, ethics and resilience in supply chain companies.

References

Al Halbusi, H., Tehseen, S. and Ramayah, T. (2017), "The impact of organisational justice on the ethical leadership under the moderating influence of perceived support: a conceptual study", *Malaysian Journal of Business and Economics*, Vol. 4 No. 1, pp. 2289-8018.

Al Issa, H.E. and Abdelsalam, M.K. (2021), "Antecedents to leadership: a CB-SEM and PLS-SEM validation", *International Journal of Sustainable Development and Planning*, Vol. 16 No. 8, pp. 1403-1414, doi: 10.18280/ijsdp.160801.

Ambrose, M.L., Arnaud, A. and Schminke, M. (2008), "Individual moral development and ethical climate: the influence of person–organisation fit on job attitudes", *Journal of Business Ethics*, Springer, Vol. 77 No. 3, pp. 323-333.

Ambrose, M.L., Rice, D.B. and Mayer, D.M. (2021), "Justice climate and workgroup outcomes: the role of coworker fair behaviour and workgroup structure", *Journal of Business Ethics*, Springer Netherlands, Vol. 172 No. 1, pp. 79-99.

Andreoli, N. and Lefkowitz, J. (2009), "Individual and organisational antecedents of misconduct in organisations", *Journal of Business Ethics*, Vol. 85 No. 3, pp. 309-332.

Arkin, R.C. and Borenstein, J. (2019), "Competing ethical frameworks mediated by moral emotions in HRI: motivations, background, and approach", *International Conference on Robot Ethics and Standards*, London, UK, pp. 29-30.

Arrami, N. and QingXiang, Y. (2021), "The role of moral identity in auditor's ethical decision making", *International Journal of Research in Business and Social Science (2147-4478)*, Vol. 10 No. 2, pp. 157-169.

Asnordin, N.A., Sundram, V.P.K. and Noranee, S. (2021), "The effect of information and communication technology and procurement towards supply chain integration", *Turkish Journal of Computer and Mathematics Education*, Karadeniz Technical University Distance Education Research and Application, Vol. 12 No. 14, pp. 1511-1520.

Audi, R. (2012), "Virtue ethics as a resource in business", *Business Ethics Quarterly*, Vol. 22 No. 2, pp. 273-291.

Bag, S., Gupta, S. and Telukdarie, A. (2018), "Exploring the relationship between unethical practices, buyer–supplier relationships and green design for sustainability", *International Journal of Sustainable Engineering*, Taylor & Francis, Vol. 11 No. 2, pp. 97-109.

Bandura, A. (1988), "Organisational applications of social cognitive theory", *Australian Journal of Management*, SAGE Publications Sage UK London, Vol. 13, pp. 275-302.

Bandura, A. (2011), "Social cognitive theory", *Handbook of Social Psychological Theories*, Vol. 2012, pp. 349-373.

Basheer, M.F., Siam, M.R.A., Awn, A.M. and Hassan, S.G. (2016), "Exploring the role of TQM and supply chain practices for firm supply performance in the presence of information technology capabilities and supply chain technology adoption: a case of textile firms in Pakistan", *Uncertain Supply Chain Management*, Vol. 7 No. 2, pp. 275-288.

Batchelor, R., Bobrowicz, A., Mackenzie, R. and Milne, A. (2012), "Challenges of ethical and legal responsibilities when technologies' uses and users change: social networking sites, decision-making capacity and dementia", *Ethics and Information Technology*, Vol. 14 No. 2, pp. 99-108.

Bessen, J. (2017), "Information technology and industry concentration", Boston University School of Law, Law & Economics Paper No. 17-41, No. 17, p. 43.

Bessen, J. (2020), "Industry concentration and information technology", *Journal of Law and Economics*, Vol. 63, doi: 10.1086/708936.

Björkdahl, J. (2020), "Strategies for digitalization in manufacturing firms", *California Management Review*, Vol. 62 No. 4, pp. 17-36.

Bolhari, A., Radfar, R., Alborzi, M., Poorebrahimi, A. and Dehghani, M. (2017), "Perceived possibility of disclosure and ethical decision making in an information technology context", *Engineering, Technology and Applied Science Research*, Vol. 7 No. 2, pp. 1567-1574.

Buntz, B. (2017), "The ordinary magic of the 21st century supply chain", Internet of Things Institute, available at: https://www.industryweek.com/technology-and-iiot/article/22006123/the-ordinary-magic-of-the-21st-century-supply-chain (accessed 28 March 2022).

Calia, C., Reid, C., Guerra, C., Oshodi, A.G., Marley, C., Amos, A., Barrera, P. and Grant, L. (2021), "Ethical challenges in the COVID-19 research context: a toolkit for supporting analysis and resolution", *Ethics and Behaviour*, Routledge, Vol. 31 No. 1, pp. 60-75.

Camilleri, P. (2012), "Teachers interpretations of the internet. An applied case study for the evaluation of technological frames of reference", *Informatics in Education*, Vol. 11 No. 2, pp. 151-167.

Carroll, N. and Conboy, K. (2020), "Normalising the 'new normal': changing tech-driven work practices under pandemic time pressure", *International Journal of Information Management*, Elsevier, Vol. 55, 102186.

Casali, G.L. (2010), "Ethical Decision making and health care managers: developing managerial profiles based on ethical frameworks and other influencing factors", Queensland University of Technology.

Casali, G.L. (2011), "Developing a multidimensional scale for ethical decision making", *Journal of Business Ethics*, Vol. 104 No. 4, pp. 485-497.

Caschili, S., Medda, F.R. and Wilson, A. (2015), "An interdependent multi-layer model: resilience of international networks", *Networks and Spatial Economics*, Vol. 15 No. 2, pp. 313-335.

Chatterjee, S., Chakraborty, S., Fulk, H.K. and Sarker, S. (2021), "Building a compassionate workplace using information technology: considerations for information systems research", *International Journal of Information Management*, Vol. 56, 102261, doi: 10.1016/j.ijinfomgt.2020.102261.

Chowdhury, P., Paul, S.K., Kaisar, S. and Moktadir, M.A. (2021), "COVID-19 pandemic related supply chain studies: a systematic review", *Transportation Research E: Logistics and Transportation Review*, Elsevier, Vol. 148, February, 102271.

Claybourn, M. (2011), "Relationships between moral disengagement, work characteristics and workplace harassment", *Journal of Business Ethics*, Vol. 100 No. 2, pp. 283-301.

Craft, J.L. (2013), "A review of the empirical ethical decision-making literature: 2004-2011", *Journal of Business Ethics*, Springer, Vol. 117 No. 2, pp. 221-259.

Criado, J.I. and de Zarate-Alcarazo, L.O. (2022), "Technological frames, CIOs, and Artificial Intelligence in public administration: a socio-cognitive exploratory study in Spanish local governments", *Government Information Quarterly*, Elsevier, Vol. 39 No. 3, 101688.

Crossan, M., Mazutis, D. and Seijts, G. (2013), "In search of virtue: the role of virtues, values and character strengths in ethical decision making", *Journal of Business Ethics*, Vol. 113 No. 4, pp. 567-581.

Crowley, F. and Bourke, J. (2018), "The influence of the manager on firm innovation in emerging economies", *International Journal of Innovation Management*, World Scientific, Vol. 22 No. 03, 1850028.

De Cremer, D. and Vandekerckhove, W. (2017), "Managing unethical behaviour in organisations: the need for a behavioral business ethics approach", *Journal of Management and Organisation*, Vol. 23 No. 3, pp. 437-455.

Deckers, T., Falk, A., Kosse, F. and Szech, N. (2021), "Homo moralis: personal characteristics, institutions, and moral decision-making", *SSRN Electronic Journal*, Vol. 066, pp. 1-25.

Demeester, K., Hopkins, A., Hodge, G. and Yustin, N. (2021), *Aligning Technology and Customer Functions in the Post-pandemic Paradigm, Russell Reynolds*, available at: https://www.russellreynolds.com/en/insights/articles/aligning-technology-customer-functions-in-the-post-pandemic-paradigm (accessed 30 March 2022).

DeVellis, R.F. (2012), "Scale development: theory and applications", in *Applied Social Research Methods*, Vol. 26, 3rd ed.

Dubey, R., Gunasekaran, A., Childe, S.J., Fosso Wamba, S., Roubaud, D. and Foropon, C. (2021), "Empirical investigation of data analytics capability and organisational flexibility as complements to supply chain resilience", *International Journal of Production Research*, Taylor & Francis Vol. 59 No. 1, pp. 110-128.

Elbe, A.-M. and Brand, R. (2016), "The effect of an ethical decision-making training on young athletes' attitudes toward doping", *Ethics and Behaviour*, Taylor & Francis, Vol. 26 No. 1, pp. 32-44.

Elm, D.R. (2019), "Cognitive moral development in ethical decision-making", *Business Ethics*, Emerald Publishing, Bingley, Vol. 3, pp. 155-177.

Engle, R.L., Lopez, E.R., Gormley, K.E., Chan, J.A., Charns, M.P. and Van Deusen Lukas, C. (2017), "What roles do middle managers play in implementation of innovative practices?", *Health Care Management Review*, Vol. 42 No. 1, pp. 14-27.

Eshkenazi, A. (2018), "Supply chain's future", available at: https://www.ascm.org/ascm-insights/scm-now-impact/supply-chains-future/ (accessed 28 March 2022).

Esmaili, M., Heidari, M. and moghaddam, V. (2021), "Process of forming unethical behaviour in business ethics: a grounded theory study in Isfahan market", *Akhlagh Religious Extension Quarterly*, Vol. 16 No. 40, pp. 127-144, doi: 10.22081/jare.2021.59146.1582.

Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Sage Publications, Vol. 18 No. 1, pp. 39-50, doi: 10.2307/3151312.

Franceli, J.C. and Zilber Turri, S.N. (2021), "Adoption factors of enabling I4.0 technologies and benefits in the supply chain", *Computer Science and Information Technology (CS & IT)*, pp. 81-96.

Frederico, G.F., Kumar, V., Garza-Reyes, J.A., Kumar, A. and Agrawal, R. (2021), "Impact of I4.0 technologies and their interoperability on performance: future pathways for supply chain resilience post-COVID-19", *International Journal of Logistics Management*, Emerald Publishing. doi: 10.1108/IJLM-03-2021-0181.

Galperin, B.L., Bennett, R.J. and Aquino, K. (2011), "Status differentiation and the protean self: a social-cognitive model of unethical behaviour in organisations", *Journal of Business Ethics*, Vol. 98 No. 3, pp. 407-424.

Gözükara, İ. and Çolakoğlu, N. (2015), "The impact of manager support and work family conflict on job satisfaction", *Business Management Dynamics*, Society of Business and Management Dynamics (SBMD), Vol. 5 No. 6, p. 13.

Guan, D., Wang, D., Hallegatte, S., Davis, S.J., Huo, J., Li, S., Bai, Y., Lei, T., Xue, Q., Coffman, D., Cheng, D., Chen, P., Liang, X., Xu, B., Lu, X., Wang, S., Hubacek, K. and Gong, P. (2020), "Global supply-chain effects of COVID-19 control measures", *Nature Human Behaviour*, Springer US, Vol. 4 No. 6, pp. 577-587.

Haddud, A. and Khare, A. (2020), "Digitalizing supply chains potential benefits and impact on lean operations", *International Journal of Lean Six Sigma*, Vol. 11 No. 4, pp. 731-765.

Haines, R. and Leonard, L.N.K. (2007), "Individual characteristics and ethical decision-making in an IT context", *Industrial Management and Data Systems*, Emerald Group Publishing, Vol. 107 No. 1, pp. 5-20.

Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., Babin, B.J. and Anderson, R.E. (2010), *Multivariate Data Analysis*, 7th ed., Pearson Prentice Hall.

Hair, J.F., Matthews, L.M., Matthews, R.L. and Sarstedt, M. (2017), "PLS-SEM or CB-SEM: updated guidelines on which method to use", *International Journal of Multivariate Data Analysis*, Inderscience Publishers, Vol. 1 No. 2, pp. 107-123.

Harapko, S. (2021), *How COVID-19 Impacted Supply Chains and What Comes Next*, Ernst & Young Global, available at: https://www.ey.com/en_gl/supply-chain/how-COVID-19-impacted-supply-chains-and-what-comes-next (accessed 28 March 2022).

Hayes, A.F. (2018), *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*, 2nd ed., The Guilford Press, New York.

Hayes, A.F., Montoya, A.K. and Rockwood, N.J. (2017), "The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling", *Australasian Marketing Journal*, Elsevier, Vol. 25 No. 1, pp. 76-81.

Henseler, J., Ringle, C.M. and Sinkovics, R.R. (2009), "The use of partial least squares path modeling in international marketing", in Sinkovics, R.R. and Ghauri, P.N. (Eds), *New Challenges to International Marketing*, Emerald Group Publishing, Vol. 20, pp. 277-319.

Herbert, W. (2010), *On Second Thought: Outsmarting Your Mind's Hard-Wired Habits*, Crown Publishers, New York.

Ho, W., Zheng, T., Yildiz, H. and Talluri, S. (2015), "Supply chain risk management: a literature review", *International Journal of Production Research*, Taylor & Francis, Vol. 53 No. 16, pp. 5031-5069.

Hohenstein, N.-O., Feisel, E., Hartmann, E. and Giunipero, L. (2015), "Research on the phenomenon of supply chain resilience: a systematic review and paths for further investigation", *International Journal of Physical Distribution and Logistics Management*, Emerald Group Publishing, Vol. 45 Nos 1/2, pp. 90-117, doi: 10.1108/IJPDLM-05-2013-0128.

Huhtala, M., Fadjukoff, P. and Kroger, J. (2021), "Managers as moral leaders: moral identity processes in the context of work", *Journal of Business Ethics*, Springer, Vol. 172 No. 4, pp. 639-652.

Hunt, S.D. and Vitell, S. (1986), "A general theory of marketing ethics", *Journal of Macromarketing*, Sage Publications Sage CA: Thousand Oaks, CA, Vol. 6 No. 1, pp. 5-16.

Hunt, S.D. and Vitell, S.J. (2006), "The general theory of marketing ethics: a revision and three questions", *Journal of Macromarketing*, Sage Publications Sage CA: Thousand Oaks, CA, Vol. 26 No. 2, pp. 143-153.

Huo, M.-L., Boxall, P. and Cheung, G.W. (2020), "How does line-manager support enhance worker wellbeing? A study in China", *The International Journal of Human Resource Management*, Taylor & Francis, Vol. 31 No. 14, pp. 1825-1843.

IBM Corp. (2016), *IBM SPSS Amos for Windows, Version 24.0*, IBM Corp., Armonk, NY.

Indian Brand Equity Foundation (IBEF) (2021), "Infrastructure", available at: https://www.ibef.org/download/Infrastructure-March-2021.pdf (accessed 29 March 2022).

Ivanov, D. (2020), "Predicting the impacts of epidemic outbreaks on global supply chains: a simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case", *Transportation Research Part E: Logistics and Transportation Review*, Elsevier, Vol. 136, March, 101922.

Ivanov, D., Dolgui, A., Sokolov, B. and Ivanova, M. (2017), "Literature review on disruption recovery in the supply chain*", *International Journal of Production Research*, Taylor & Francis, Vol. 55 No. 20, pp. 6158-6174.

Jia, X. (2019), "Essays on ethics in supply chain management", Doctoral dissertation, Texas A M University, available at: https://oaktrust.library.tamu.edu/handle/1969.1/184952.

Jones, T.M. (1991), "Ethical decision making by individuals in organizations: an issue-contingent model", *Academy of Management Review*, Vol. 16 No. 2, pp. 366-395.

Jordan, P.J. and Troth, A.C. (2020), "Common method bias in applied settings: the dilemma of researching in organisations", *Australian Journal of Management*, Vol. 45 No. 1, pp. 3-14.

Kähkönen, A.K., Evangelista, P., Hallikas, J., Immonen, M. and Lintukangas, K. (2021), "COVID-19 as a trigger for dynamic capability development and supply chain resilience improvement", *International Journal of Production Research*, Taylor & Francis Online, pp. 1-20. doi: 10.1080/00207543.2021.2009588.

Kanjanapathy, M. and Ramakrishnan, S. (2018), "The influence of knowledge, skills, attitude and organisation commitment using ethical climate as A moderator towards ethical decision making of accountants in Malaysia: a conceptual paper", *International Journal of Engineering and Technology*, Vol. 7, p. 1155.

Kaplan, S. and Tripsas, M. (2008), "Thinking about technology: applying a cognitive lens to technical change", *Research Policy*, Vol. 37 No. 5, pp. 790-805.

Katsaliaki, K., Galetsi, P. and Kumar, S. (2021), "Supply chain disruptions and resilience: a major review and future research agenda", *Annals of Operations Research*, Springer US. doi:10.1007/s10479-020-03912-1.

Khalid, K., Wan Omar, M., Syed Agil, S.O. and Khalid, K. (2012), "The effects of individual characteristics on ethical decision making in SMEs", *International Journal of Academic Research*, Vol. 5 No. 1, pp. 18-22.

Khan, H.ur R., Usman, B., Zaman, K., Nassani, A.A., Haffar, M. and Muneer, G. (2022), "The impact of carbon pricing, climate financing, and financial literacy on COVID-19 cases: go-for-green healthcare policies", *Environmental Science and Pollution Research*, Springer Berlin Heidelberg, Vol. 29 No. 24, pp. 35884-35896.

Kim, T.W., Lee, H., Kim, M.Y., Kim, S.A. and Duhachek, A. (2022), "AI increases unethical consumer behaviour due to reduced anticipatory guilt", *Journal of the Academy of Marketing Science*, Springer. doi: 10.1007/s11747-021-00832-9.

Kumar, M.S., Raut, D.R.D., Narwane, D.V.S. and Narkhede, D.B.E. (2020), "Applications of industry 4.0 to overcome the COVID-19 operational challenges", *Diabetes and Metabolic Syndrome*, Diabetes India. Published by Elsevier, Vol. 14 No. 5, pp. 1283-1289.

Kutnjak, A., Pihiri, I. and Furjan, M.T. (2019), "Digital transformation case studies across industries – literature review", *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1293-1298.

Laaper, S., Yauch, G., Wellener, P. and Robinson, R. (2018), "Embracing a digital future: how manufacturers can unlock the transformative benefits of digital supply networks", available at: https://www2.deloitte.com/us/en/insights/focus/industry-4-0/digital-supply-network-transformation-study.html (accessed 28 March 2022).

Lehnert, K., Park, Y.hwal and Singh, N. (2015), "Research note and review of the empirical ethical decision-making literature: boundary conditions and extensions", *Journal of Business Ethics*, Springer Netherlands, Vol. 129 No. 1, pp. 195-219.

Lin, X., Clay, P.F., Hajli, N. and Dadgar, M. (2018), "Investigating the impacts of organisational factors on employees' unethical behaviour within organisation in the context of Chinese firms", *Journal of Business Ethics*, Springer, Vol. 150 No. 3, pp. 779-791.

Lind, G. (1998), "An introduction to the moral judgment test (MJT)", Unpublished manuscript, University of Konstanz, Konstanz, available at: http://www.uni-konstanz.de/ag-moral/pdf/mjt-introduction.pdf.

Lindell, M.K. and Whitney, D.J. (2001), "Accounting for common method variance in cross-sectional research designs", *Journal of Applied Psychology*, American Psychological Association, Vol. 86 No. 1, p. 114.

Liu, K.P. and Chiu, W. (2021), "Supply Chain 4.0: the impact of supply chain digitalization and integration on firm performance", *Asian Journal of Business Ethics*, Vol. 10 No. 2, pp. 371-389.

Lopes de Sousa Jabbour, A.B., Chiappetta Jabbour, C.J., Hingley, M., Vilalta-Perdomo, E.L., Ramsden, G. and Twigg, D. (2020), "Sustainability of supply chains in the wake of the coronavirus (COVID-19/SARS-CoV-2) pandemic: lessons and trends", *Modern Supply Chain Research and Applications*, Vol. 2 No. 3, pp. 117-122.

López-Fernández, A.M. and Sánchez-Rosales, A. (2022), "Unethical supply chains delaying recovery: analyzing pre and mid COVID-19 conditions", *Business Recovery in Emerging Markets*, Springer, pp. 201-222.

Luca Casali, G. and Perano, M. (2021), "Forty years of research on factors influencing ethical decision making: establishing a future research agenda", *Journal of Business Research*, Elsevier, Vol. 132, February, pp. 614-630.

Mâta, L., Poenaru, A.-G. and Boghian, I. (2021), "Current issues of ethical use of information technology from the perspective of university teachers", *Ethical Use of Information Technology in Higher Education*, Springer Nature, pp. 163-179.

Malhotra, N.K. (2016), *Essentials of Marketing Research: A Hands-On Orientation*, 7th ed., Pearson Education.

Malloy, D.C. and Agarwal, J. (2010), "Ethical climate in government and nonprofit sectors: public policy implications for service delivery", *Journal of Business Ethics*, Springer, Vol. 94 No. 1, pp. 3-21.

Martinez, L.F. and Jaeger, D.S. (2016), "Ethical decision making in counterfeit purchase situations: the influence of moral awareness and moral emotions on moral judgment and purchase intentions", *Journal of Consumer Marketing*, Vol. 33 No. 3, pp. 213-223.

Matyas, D. and Pelling, M. (2015), "Positioning resilience for 2015: the role of resistance, incremental adjustment and transformation in disaster risk management policy", *Disasters*, Wiley Online Library, Vol. 39 No. s1, pp. s1-s18.

McCullough, M.E., Emmons, R.A. and Tsang, J.-A. (2002), "The grateful disposition: a conceptual and empirical topography", *Journal of Personality and Social Psychology*, American Psychological Association, Vol. 82 No. 1, p. 112.

McKibbin, W.J. and Fernando, R. (2020), "The global macroeconomic impacts of COVID-19: seven scenarios", CAMA Working Paper No. 19/2020, *SSRN Electronic Journal*, doi: 10.2139/ssrn.3547729.

Melnyk, S.A., Closs, D.J., Griffis, S.E., Zobel, C.W. and Macdonald, J.R. (2014), "Understanding supply chain resilience", *Supply Chain Management Review*, Vol. 18 No. 1, pp. 34-41.

Mishra, A.N. and Agarwal, R. (2010), "Technological frames, organisational capabilities, and IT use: an empirical investigation of electronic procurement", *Information Systems Research*, Vol. 21 No. 2, pp. 249-270.

Moore, G. (2005), "Corporate character: modern virtue ethics and the virtuous corporation", *Business Ethics Quarterly*, Vol. 15 No. 4, pp. 659-685, doi: 10.5840/beq200515446.

Moore, C., Detert, J.R., Klebe Treviño, L., Baker, V.L. and Mayer, D.M. (2012), "Why employees do bad things: moral disengagement and unethical organisational behaviour", *Personnel Psychology*, Vol. 65 No. 1, pp. 1-48.

Nasiri, M., Ukko, J., Saunila, M. and Rantala, T. (2020), *Managing the Digital Supply Chain: the Role of Smart Technologies*", *Technovation*, Elsevier, March, pp. 96-97, 102121.

Newman, D. (2018), "How IoT will impact the supply chain", *Forbes*, available at: https://www.forbes.com/sites/danielnewman/2018/01/09/how-iot-willimpact-the-supply-chain (accessed 25 March 2022).

Odole, E.B. (2018), "Factors that encourage unethical practices by organisational leaders in Nigeria", ProQuest Dissertations and Theses, p. 203.

Oh, I. and Shim, D. (2020), "It adoption and sustainable growth of firms in different industries—are the benefits still expected?", *Sustainability (Switzerland)*, Vol. 12 No. 22, pp. 1-29.

Olesen, K. (2014), "Technological Frames", *SAGE Open*, Vol. 4 No. 1, pp. 1-15, doi: 10.1177/2158244014526720.

Olesen, K., Narayan, A.K. and Ramachandra, S. (2013), "Technological frames: influence of group frames (ingluence)", *Problems and Perspectives in Management*, Vol. 11 No. 1, pp. 81-93.

Orlikowski, W.J. and Gash, D.C. (1994), "Technological frames: making sense of information technology in organisations", *ACM Transactions on Information Systems (TOIS)*, Vol. 12 No. 2, pp. 174-207, doi: 10.1145/196734.196745.

Otulugbu, D. and Ogunobo, M. (2022), "Ethical issues and information communication technology (ICT) use in the new era", *International Journal of Knowledge Content Development and Technology*, Vol. 11 No. 30.

O'Driscoll, M.P., Brough, P., Timms, C. and Sawang, S. (2010), "Engagement with information and communication technology and psychological well-being", *New Developments in Theoretical and Conceptual Approaches to Job Stress*, Emerald Group Publishing, Vol. 8, pp. 269-316.

Park, S., Kang, H.S., (Theresa) and Kim, E.J. (2018), "The role of supervisor support on employees' training and job performance: an empirical study", *European Journal of Training and Development*, Vol. 42 Nos 1-2, pp. 57-74.

Parungao, A. (2020), "The future of remote work after COVID-19: 3 common predictions", available at: https://www.ekoapp.com/blog/the-future-of-remote-work-after-covid-19-3-common-predictions (accessed 25 March 2022).

Paul, S.K., Moktadir, M.A. and Ahsan, K. (2021), "Key supply chain strategies for the post-COVID-19 era: implications for resilience and sustainability", *International Journal of Logistics Management*, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/IJLM-04-2021-0238.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879-903, doi: 10.1037/0021-9010.88.5.879.

Provdanov, C.C. and De Freitas, E.C. (2013), *Metodologia Do Trabalho Científico: Métodos e Técnicas Da Pesquisa e Do Trabalho Acadêmico*, Feevale, Novo Hamburgo, doi: 10.1017/CBO9781107415324.004.

PwC (2021), "COVID-19: operations and supply chain disruption", available at: https://www.pwc.com/us/en/library/covid-19/supply-chain.html (accessed 28 March 2022).

Rajesh, R. (2022), "A novel advanced grey incidence analysis for investigating the level of resilience in supply chains", *Annals of Operations Research*, Springer US, Vol. 308, doi: 10.1007/s10479-020-03641-5.

Reidenbach, R.E. and Robin, D.P. (1990), "Toward the development of a multidimensional scale for improving evaluations of business ethics", *Journal of Business Ethics*, Springer, Vol. 9 No. 8, pp. 639-653.

Resende, M.M. and Porto, J.B. (2020), "Brazilian jeitinho and moral identity: impact on ethical behaviour at work", *Trends in Psychology*, Springer, Vol. 28 No. 1, pp. 148-164.

Rest, J.R. (1979), *Development in Judging Moral Issues*, University of Minnesota Press, Minneapolis, MN.

Roh, J., Tokar, T., Swink, M. and Williams, B. (2022), "Supply chain resilience to low-/high-impact disruptions: the influence of absorptive capacity", *International Journal of Logistics Management*, Vol. 33 No. 1, pp. 214-238.

Saberi, S., Kouhizadeh, M., Sarkis, J. and Shen, L. (2019), "Blockchain technology and its relationships to sustainable supply chain management", *International Journal of Production Research*, Vol. 57 No. 7, pp. 2117-2135.

Sahoo, M. (2019), "Structural equation modeling: threshold criteria for assessing model fit", *Methodological Issues in Management Research: Advances, Challenges, and the Way Ahead*, Emerald Publishing, Bingley, pp. 269-276.

Sanders, N. and Morgan, S. (2020), "Digital supply chain transformation: visualizing the possibilities", *Logistics Management*, Vol. 58 No. 3, pp. 42-53.

Sarkar, P., Mohamed Ismail, M.W. and Tkachev, T. (2022), "Bridging the supply chain resilience research and practice gaps: pre and post COVID-19 perspectives", *Journal of Global Operations and Strategic Sourcing*, Vol. ahead-of-print, doi: 10.1108/jgoss-09-2021-0082.

Sarkis, J. (2021), "Supply chain sustainability: learning from the COVID-19 pandemic", *International Journal of Operations and Production Management*, Vol. 41 No. 1, pp. 63-73.

Sashkin, M., Rosenbach, W.E. and Sashkin, M.G. (1997), "Development of the power need and its expression in leadership and management with a focus on leader-follower relations", *Proceedings of the Twelfth Scientific Meeting of the AK Rice Institute*, AK Rice Institute, Leadership as Legacy, Jupiter, FL.

Scholten, K. and Schilder, S. (2015), "The role of collaboration in supply chain resilience", *Supply Chain Management*, Vol. 20 No. 4, pp. 471-484.

Schwartz, M.S. (2016), "Ethical decision-making theory: an integrated approach", *Journal of Business Ethics*, Springer Netherlands, Vol. 139 No. 4, pp. 755-776.

Schwarz, A., Rizzuto, T., Carraher-Wolverton, C., Roldán, J.L. and Barrera-Barrera, R. (2017), "Examining the impact and detection of the 'urban legend' of common method bias", *Data Base for Advances in Information Systems*, Vol. 48 No. 1, pp. 93-119.

Shanahan, K.J. and Hyman, M.R. (2003), "The development of a virtue ethics scale", *Journal of Business Ethics*, Vol. 42 No. 2, pp. 197-208.

Shishodia, A., Sharma, R., Rajesh, R. and Munim, Z.H. (2021), "Supply chain resilience: a review, conceptual framework and future research", *International Journal of Logistics Management*, Vol. ahead-of-print, doi: 10.1108/IJLM-03-2021-0169.

Shrestha, P. (2021), "Use of technology and its management issues in Nepalese industries and businesses", *Nepalese Journal of Management Research*, Vol. 1, pp. 9-14.

Simangunsong, E., Hendry, L.C. and Stevenson, M. (2016), "Managing supply chain uncertainty with emerging ethical issues", *International Journal of Operations and Production Management*, Vol. 36 No. 10, pp. 1272-1307.

Smolinski, P.R., Kowalik, J. and Winiarski, J. (2022) "Diversity of students' unethical behaviors in online learning amid COVID-19 pandemic: an exploratory analysis BT - information systems", in Themistocleous, M. and Papadaki, M. (Eds), Springer International Publishing, Cham, pp. 551-566.

Sönmez, B. and Yıldırım, A. (2019), "The mediating role of autonomy in the effect of pro-innovation climate and supervisor supportiveness on innovative behaviour of nurses", *European Journal of Innovation Management*, Vol. 22 No. 1, pp. 41-58.

Spieske, A. and Birkel, H. (2021), "Improving supply chain resilience through industry 4.0: a systematic literature review under the impressions of the COVID-19 pandemic", *Computers and Industrial Engineering*, Elsevier, June, Vol. 158, 107452.

Spieth, P., Röth, T., Clauss, T. and Klos, C. (2021), "Technological frames in the digital age: theory, measurement instrument, and future research areas", *Journal of Management Studies*, Vol. 58 No. 7, pp. 1962-1993.

Sriyanto, S., Lodhi, M.S., Salamun, H., Sardin, S., Pasani, C.F., Muneer, G. and Zaman, K. (2022), "The role of healthcare supply chain management in the wake of COVID-19 pandemic: hot off the press", *Foresight*, Vol. 24 Nos 3/4, pp. 429-444.

Tabachnick, B.G. and Fidell, L.S. (2007), "Using multivariate statistics, PsycCRITIQUES", doi:10.1037/022267.

Tjahjono, B., Esplugues, C., Ares, E. and Pelaez, G. (2017), "What does industry 4.0 mean to supply chain?", *Procedia Manufacturing*, Elsevier B.V., Vol. 13, pp. 1175-1182.

Tyre, M.J. and Orlikowski, W.J. (1994), "Windows of opportunity: temporal patterns of technological adaptation in organisations", *Organisation Science*, INFORMS, Vol. 5 No. 1, pp. 98-118.

Ünal, A.F., Warren, D.E. and Chen, C.C. (2012), "The normative foundations of unethical supervision in organisations", *Journal of Business Ethics*, Springer, Vol. 107 No. 1, pp. 5-19.

van Hoek, R. (2019), "Exploring blockchain implementation in the supply chain: learning from pioneers and RFID research", *International Journal of Operations and Production Management*, Vol. 39, pp. 829-859.

Van Hulst, M. and Yanow, D. (2016), "From policy 'frames' to 'framing' theorizing a more dynamic, political approach", *The American Review of Public Administration*, SAGE Publications Sage CA: Los Angeles, CA, Vol. 46 No. 1, pp. 92-112.

Verma, S. and Garg, N. (2022), "Exploring intergenerational differences in technology-oriented ethical behaviour", *Kybernetes*, Emerald Publishing, Vol. ahead-of-print No. ahead-of-print, doi: 10.1108/K-09-2021-0897.

Vollmer, M. (2018), "Artificial intelligence: what's now and next in IoT-driven supply chain innovation. Silicon semiconductor", available at: https://siliconsemiconductor.net/article/104738/Artificial_Intelligence_What (accessed 25 March 2022).

Wang, Z., Xing, L., Xu, H. and Hannah, S.T. (2021), "Not all followers socially learn from ethical leaders: the roles of followers' moral identity and leader identification in the ethical leadership process", *Journal of Business Ethics*, Springer, Vol. 170 No. 3, pp. 449-469.

Wieland, A. and Durach, C.F. (2021), "Two perspectives on supply chain resilience", *Journal of Business Logistics*, Vol. 42 No. 3, pp. 315-322.

Winkielman, P., Knutson, B., Paulus, M. and Trujillo, J.L. (2007) "Affective influence on judgments and decisions: moving towards core mechanisms", *Review of General Psychology*, SAGE Publications Sage CA: Los Angeles, CA, Vol. 11 No. 2, pp. 179-192.

Wittmer, D.P. (2019), "Ethical decision-making", *Handbook of Administrative Ethics*, Routledge, pp. 481-507.

Yun, G., Ebrahimpour, M., Bandyopadhyay, P. and Withers, B. (2020), "Internal and vendor employees' unethical behaviors in the supply chain: the case of India", *Benchmarking: An International Journal*, Vol. 27 No. 1, pp. 59-80.

Zardeckaite-Matulaitiene, K., Endriulaitiene, A. and Naujokaitiene, J. (2014), *Employees' Attitudes toward Ethical Behaviour at Work: Impact of Their Individual Characteristics*, Brill, Leiden, pp. 81-96.

Zeni, T.A., Buckley, M.R., Mumford, M.D. and Griffith, J.A. (2016), "Making 'sense' of ethical decision making", *Leadership Quarterly*, Vol. 27 No. 6, pp. 838-855.

Zhang, A.Y., Tsui, A.S., Song, L.J., Li, C. and Jia, L. (2008), "How do I trust thee? The employee-organisation relationship, supervisory support, and middle manager trust in the organisation", *Human Resource Management Published in Cooperation with the School of Business Administration, The University of Michigan and in Alliance with the Society of Human Resources Management*, Wiley Online Library, Vol. 47 No. 1 pp. 111-132.

Zollo, L., Pellegrini, M.M. and Ciappei, C. (2017), "What sparks ethical decision making? The interplay between moral intuition and moral reasoning: lessons from the scholastic doctrine", *Journal of Business Ethics*, Springer, Vol. 145 No. 4, pp. 681-700.

Zuber, F. (2015), "Spread of unethical behaviour in organisations: a dynamic social network perspective", *Journal of Business Ethics*, Springer, Vol. 131 No. 1, pp. 151-172.

Appendix
Instruments used

A. Techno-ethical orientation (Verma and Garg, 2022)

Instructions – Please rate your response according to the scale provided below:
1 – highly unethical; 2 – unethical; 3 – neutral; 4 – ethical; 5 – highly ethical.

(1) Reading and forwarding emails of others without their consent
(2) While at work, browsing the organisation Internet for personal needs
(3) Sending files with viruses to others intentionally
(4) Sharing confidential information with others on social media/online platforms
(5) Recording meetings without prior permission from employer/supervisor
(6) Taking pictures of colleagues in the premises without their awareness
(7) Copying software from the organisation for personal use
(8) Printing documents at the organisation for personal use
(9) Not giving due credits to someone for providing project-related material
(10) Discussing negative aspects about the colleagues/organisation with others on Internet
(11) Using organisation's Internet for recreational purposes
(12) Using instant messaging apps during office hours in the organisation
(13) Downloading files at the organisation from the Internet for personal use
(14) Tracking the data from confidential sources to threaten or defame the other person

B. Ethical decision-making (Casali, 2011)

Instructions – When fulfilling the requirements of your position in your organisation, please indicate the importance of the followings in your decision-making process.

Rate your level of agreement/disagreement according to the scale provided below:
1 – Extremely important; 2 – Very important; 3 – Fairly important; 4 – Not very important; 5 – Not important at all.

(1) Providing the highest economic return (profit) for the organisation
(2) Minimizing costs for the organisation
(3) Protecting the reputation of the organisation
(4) Optimizing resources of the unit/department
(5) Attaining organisational yearly budgets (short-term)
(6) Being in line with the organisational mission
(7) Generating the greatest overall benefits for the department/organisation
(8) Not harming the clients/vendors
(9) Respecting organisational rules and regulations that have been created for the greatest benefit of all stakeholders
(10) Obeying the law (state and central)
(11) Creating the greatest overall benefit for the local community
(12) Creating the greatest overall benefit for the wider community

(13) Being most in line with your core personal values
(14) Being most in line with the person you want to be
(15) Respecting dignity of those affected by the decision
(16) Being able to empathize with clients/vendors
(17) Acting openly when making decision
(18) Making "care for the affected" is paramount in determining decision alternatives
(19) Giving the opportunity to all affected parties or their representatives to have input into the decision-making process
(20) Treating others as you want others to treat you
(21) Treat people as ends not as means
(22) Ensuring that confidentiality is maintained at all times
(23) Maintaining a fair process at all times
(24) Ensuring that the organisation "duty of care" is maintained at all times

C. Technological Frames (Spieth *et al.*, 2021)
Rate your level of agreement/disagreement according to the scale provided as below:
1 – strongly disagree; 2 – disagree; 3 – somewhat disagree; 4 – neither agree nor disagree; 5 – somewhat agree; 6 – agree; 7 – strongly agree.

Personal Attitude

(1) My attitude towards technologies is positive
(2) I have high expectations of technologies
(3) Technologies are an important part of my life
(4) I regularly try to obtain information about technologies

Application Value

(1) Technologies could facilitate the coordination of my work tasks
(2) Technologies make my work more flexible
(3) Technologies reduce the possibility of making mistakes in my work
(4) Technologies increase the effectiveness of my work steps

Organisational Influence

(1) My colleagues remind me to use technologies in my job
(2) My colleagues regularly recommend technologies to me
(3) My colleagues demand that I use technologies for my job
(4) My colleagues help me use technologies for my job

Supervisor Influence

(1) My supervisor is willing to integrate technologies into the firm

(2) My supervisor requests that I use technologies

(3) My supervisor regularly speaks about technologies

(4) My supervisor is an expert in the handling of technologies

Industry Influence

(1) Our competitors demand the use of technologies

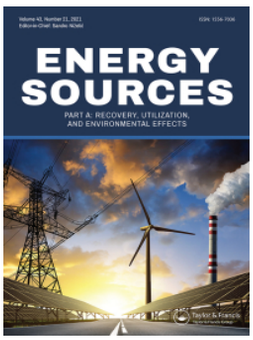
(2) Our competitors successfully use technologies

(3) Our customers demand the use of technologies

(4) Our suppliers demand the use of technologies

Corresponding author
Naval Garg can be contacted at: naval.garg@dtu.ac.in

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**



Taylor & Francis
Taylor & Francis Group

Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ueso20

Circularity issues and blockchain technology in the auto industry

Syed Wasiul Hasan Rizvi, Saurabh Agrawal & Qasim Murtaza

To cite this article: Syed Wasiul Hasan Rizvi, Saurabh Agrawal & Qasim Murtaza (2022) Circularity issues and blockchain technology in the auto industry, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 44:3, 7132-7144, DOI: 10.1080/15567036.2022.2107119

To link to this article: https://doi.org/10.1080/15567036.2022.2107119

Published online: 31 Jul 2022.

Submit your article to this journal

Article views: 27

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ueso20

ENERGY SOURCES, PART A: RECOVERY, UTILIZATION, AND ENVIRONMENTAL EFFECTS
2022, VOL. 44, NO. 3, 7132–7144
https://doi.org/10.1080/15567036.2022.2107119

Taylor & Francis
Taylor & Francis Group

Check for updates

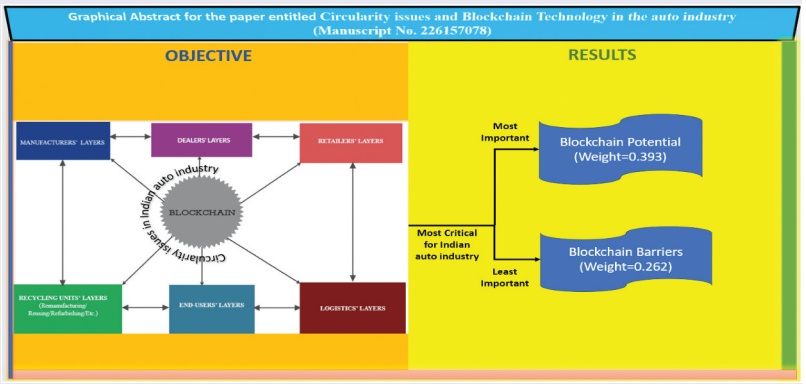
Circularity issues and blockchain technology in the auto industry

Syed Wasiul Hasan Rizvi[a], Saurabh Agrawal[b], and Qasim Murtaza[c]

[a]Department of Mechanical Engineering (DME), Delhi Technological University (DTU), Delhi, India; [b]Delhi School of Management, DTU, Delhi, India; [c]Department of Mechanical Engineering, DTU, Delhi, India

ABSTRACT

The integration of circular economy and Industry 4.0 components like blockchain technology results in a faster induction of 'circularity' and leads to new strategies on resource utilization, making the world more sustainable. It is a 'key-enabler' in meeting tomorrow's circularity challenges. Literature revealed research gaps in a yet unexplored area of how 'circularity' building gets impacted by blockchain, especially in the Indian auto sector. Moreover, the nonavailability of its evaluation tool also necessitated present research. The main objective was to explore, from the auto stakeholders' perception viewpoints, the potential of the application of blockchain technology with special reference to the issue of 'sustainability and circularity' in the Indian context. The study involved a literature search, building a new circular economy-based blockchain model for the auto industry, and, finally, a case study in which the hypotheses about circularity and blockchain technology nexus impacts on the Indian auto industry were structured and tested through a research instrument, designed and sent to 45 auto stakeholders for evaluation of the said nexus. The data analyzed through the Best-Worst method revealed the 'potential' (weight: 0.393) and 'application' (weight: 0.262) impacts on CE are the most important yardsticks. The proposed model appears to be more reliable and secure and might help the managers of auto firms in forging new circularity-based strategies to exploit I4.0 efficiently.



ARTICLE HISTORY
Received 29 April 2022
Revised 20 July 2022
Accepted 21 July 2022

KEYWORDS
Circular economy; industry 4.0; blockchain technology; auto industry; best worst method

Introduction

In today's smart technology-based world, digitization has led to the fourth Industrial revolution (I 4.0) while neither academia nor industry provides enough guidelines for the evaluation of the evolving digitized industrial culture based on IT tools (Gigauri 2021). Around 1.5 trillion USD worth would be

CONTACT Syed Wasiul Hasan Rizvi wasiulhasan@yahoo.com Department of Mechanical Engineering (DME), , 302, N-6, Jamianagar, New Delhi, India

© 2022 Taylor & Francis Group, LLC

yielded by the digitization of the logistics processes by 2025 (World Economic Forum White Paper 2016). I4.0 that relies on digitization enables the development of the circular economy (CE) at a faster pace. CE keeps the resources always in 'circulation' and helps in sustainability development. Definition-wise, CE is 'an economy that is restorative and regenerative by design which aims to keep products, components, and materials at their highest utility and value at all times, distinguishing between technical and biological cycles.' The I4.0 and CE interface, emerging as 'an area of focus' (Okorie et al. 2018), enables the evolution of strategies for new natural resources, improves productivity, and results in sustainability. The applications of this interface are spread over many areas like energy management, procurement, scheduling, and quality management in manufacturing, firefighting, hospitals, construction, marketing, etc. (Awan, Sroufe, and Shahbaz 2021). In light of the absence of the industrial-digitization evaluation guidelines, as stated above, recent developments in the field indicate an urgent need for evaluation of digital technology in the implementation of 'circularity' in industrial houses, especially, when it is noted that the related areas 'are yet severely under researched' (Alhawari et al. 2021). Particularly, in the auto industry, in this context, there appears to be a gap in research. Also, an understanding of the interest of stakeholders like suppliers and customers, in the context of CE initiatives, is needed and presently there is a lack of research on the topic (Awan, Sroufe, and Shahbaz 2021). Thus, the present research, perhaps, for the first time, contributes to the knowledge of the auto industry stakeholders on the I4.0 and CE nexus. The main objective of the present research was to explore, from the auto stakeholders' viewpoints, the potential of the application of blockchain technology (BT), in the auto industry with special reference to the issue of 'sustainability and circularity' in the Indian context. The novelty of this study becomes more apparent when it is noted that the need of the time is to explore the relationship between CE initiatives and the stakeholders' perceptions and expectations (Govindan and Hasanagic 2018). Such studies are also needed because the auto industry in India and elsewhere too plays a critical role in industrial development and employment generation. The intent of this work lies in the variety of perceptions obtained from different kinds of auto stakeholders considered, thereby enriching the literature on the study of the CE-BT in the auto industry. Therefore, the present study would go a long way in policy planning in the field of CE and BT in developing nations of the world. The remaining parts of the contents of the paper are structured as follows.

Materials and methods

Literature search

The methodology adopted for the present literature search was 'content analysis' (Rizvi, Agrawal, and Murtaza 2021) as summarized under the following heads:

BT as a tool of digitization

BT, among the top 10 technologies (Rayome 2019) comprises 'blocks of data' connected through 'cryptographic pointers' in its 'ledger' It has a numerous potential providing complete transparency and trust and is equipped with many basic features. It is a powerful tool for designing future smart systems and cities generating social sustainability China reported huge BT applications including initiatives like Blockchain Network (BCN) and digital yuan and makes supply chain management highly acceptable (Kshetri 2018). Though it has yet not reached the desired scale of BT maturity (Wang et al. 2019), BT application areas are getting widely spread globally reaching $2.3b market (Bumblauskas et al. 2019). In general, digital tools' applications result in smart companies and effectuate digital sustainability. Despite theoretical support for BT, varieties of barriers and challenges in the path of BT applications have also been reported by researchers. Thus, under the umbrella of I4.0 developments, varieties of technologies including the emerging BT have evolved. Even in the most precious and universally needed sector, energy, BT helped the 'decentralized' energy management in terms of distribution, trade, and finance-related functions (Rejeb et al. 2022). However, as observed by Gigauri (2021), studies pertaining to their evaluation appear to be missing in the literature.

CE under the impact of BT

CE means differently to different researchers as demonstrated by Kirchherr's work of 114 definitions (Kirchherr, Reike, and Hekkert 2017) and it is 'characterized as an economy that is restorative and regenerative by design that aims to keep products, components, and materials at their highest utility and value at all times, distinguishing between technical and biological cycles' (MacArthur and Waughray 2016). Industry 4.0 is ready to accept the impact of the sustainable industrial future. Among Fortune 100 companies, there were 82 which are employing BT (Schinckus, 2020; Cheng et al. 2021). However, the employment of BT appears to be yet only in the conceptual and pilot stages (Zhang and Guin 2020). Moreover, the evaluation of such newer technologies like BT is difficult due to the IT-based complexities and thus evaluation process presents another emerging topic of research for tomorrow. One major problem is high energy consumption and pollution costs due to the emission problems for which Rashid Khan et al. (2021) suggested framing new policies to tackle the problem more effectively. Chen et al. (2020) also studied the impact of technological innovations on the emission problem. To sum up, despite the high energy cost of BT, its application is expanding rapidly. Within the CE framework, BT enables its application faster thereby indicating the CE-BT interface to be a promising area of research in the days ahead.

CE-BT implications for the auto industry

Today's transportation systems through 'smart cars' are getting revolutionized in the days ahead. BT-based VID (Vehicle IDentity) appears to be a promising tool for the future. The CE strategy based on recycling in the auto industry would have its short-loop constituents as varieties of sensors involving Big Data technology (Awan, Sroufe, and Shahbaz 2021; Awan et al, 2021a; Awan et al, 2021b). To sum up, the literature revealed that the design process of smart cars is advancing day by day, and the induction of smartness in products results in the evolution of newer models of business and so is the case with auto manufacturing also. In addition, in the context of the digitized CE-auto framework, it was found that the study of the stakeholders' role emerged to be a 'prerequisite for circular manufacturing' (Awan, Sroufe, and Shahbaz 2021).

The issues to be addressed, the knowledge gaps, and the research implications

The above-presented literature search indicated that exploring issues like circularity and sustainability in the auto industry within the framework of CE-BT is the need of the hour. This is due to the fast-spreading use of the digitized industrial culture and rapidly growing technology of 'smartness' in the auto industry at the global as well as the national levels. As induction of BT is a recent development in manufacturing, studies pertaining to CE-BT and that too in the auto industry involving the perception of its stakeholders appear to be either few or none, especially in the Indian context. So far as BT is concerned, its 'adoption' is still in the naïve stage (Zhao et al. 2019), particularly in India (Yadav et al. 2020). Thus, it might be concluded that there exists a knowledge gap in terms of the availability of the data pertaining to the auto stakeholders' perception. Such data might be helpful and needed by the future policy planners for the success of digitized CE or the CE-BT-based auto manufacturing, particularly, in the Indian context. Such studies would have long-term implications for the future auto manufacturers under the upcoming and ongoing impact of the I4.0 in the industrial world. An understanding of I4.0 components that include the BT also would contribute to evolving a sustainable society and help in laying down policies on I4.0 in the future. As of now, Kamble et al. (2021) observed that 'there is a significant lack of knowledge' on the sustainability-I4.0 relationship. Such knowledge gaps are being addressed by the present work.

Proposed operational CE-BT model for the auto industry

CE closes the open-loop by way of induction of another connecting box called 'recycling.' In terms of BT, 'tasks' are 'transactions' while their 'managers' are 'agents' and the same would be employed in the proposed model also. The BT networks would involve 'decentralized' constituents under different

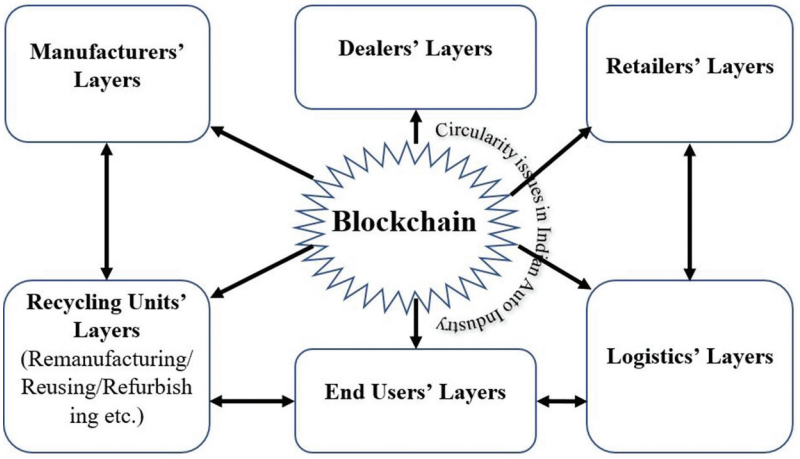


Figure 1. The CE-BT model structure proposed for the BT-based CE application in the auto industry.

'agents' who need not be 'trustworthy.' The characteristics of 'reliability' and 'security' are inducted through the 'cryptographic' nature of the BT system. All the transactions are recorded on multilayered ledgers by different users of the CE-BT system. The induction of blockchain in the new model of the auto supply chain would enable all the users of the auto network situated in their respective decentralized locations to record and read their respective transactions in their respective ledgers. All the activities of the CE-BT users are coordinated by the 'blockchain' at the center whereas its layers constitute the satellites surrounding the blockchain as is shown in Figure 1.

Evaluation of Indian auto stakeholders' perception of CE-BT use: a case study

The following sections explain the conceptual framework of the present study along with the research and data methodology employed to test the hypotheses structured as given below.

Conceptual framework and hypotheses development

The procedure of the 'Evaluation of the Perception of Auto Stakeholders' (EPASH) on CE-BT is presented conceptually as an MCDM problem as shown in Table 1 in which three primary criteria to measure EPASH in CE-BT study were BT potential (BTP), BT barriers (BTB) and BT application (BTA) with their 5 respective sub-criteria as BTP1 to BTP5; BTB1to BTB5, and BTA1 to BTA5. To test the CE-BT impacts following hypotheses were structured:

Table 1

H1: All the considered factors related to 'BTP,' 'BTB,' and 'BTA,' within the CE framework have the same influence on the stakeholders' perceived decision-making process.

H2: All the considered three (manufacturers, dealers, and end-users) Indian auto stakeholders' perceptions about CE-BT impacts weigh equally in terms of the investigated factors.

Data obtained from the stakeholders lead to sound analysis and good results (Zhang 2019).

Table 1. Measures of EPASH on CE-BT nexus impacts as adopted from different sources.

Main Criteria/ Sub-Criteria	Sources	Main Criteria/Sub-Criteria	Sources
BT Potential (BTP)/ BTP1 to BTP5 *BT Barriers (BTB/* BTB1 to BTB5	Upadhyay et al, 2021; Dalenogare et al. 2018;	*BT Applications (BTA)/* BTA1 to BTA5	Gerger 2021

Research and data methodology

The data sets utilized are the responses of the three auto stakeholders in terms of their perception of CE-BT impacts on the auto industry (expressed on a Likert scale of 1(most important) to 5 (least important)) obtained through the questionnaire (Appendix-A). The data sets included the queries pertaining to CE-BT impacts and the response entry cells. Variables were the decision criteria (3 in number:) and their sub-criteria (5 in number). Thus, there are 3 × 5 = 15 variables considered. The indicators, the base of the analysis, are the respondents' scores processed through BWM-based weighting procedures.

Research instrument employed and data sets. To get the measures of the EPASH two sets of data were needed: (a) Optimal WEIGHT for the criteria, (w_1*, w_2*, . . .,w_n*) and (b) EPASH SCORES (on the various criteria). To determine weights, five IT experts' opinions were considered, and to determine scores, data were collected through the research instrument (Appendix-A) employed for surveying the auto stakeholders, spread over 30 users, 5 dealers and 10 manufacturers, mostly from north India.

Respondents' profiles. In surveying, the auto companies targeted were those having experience in the CE and BT. Respondents, participating voluntarily desired their data to be kept confidential. Their demographic profiles are presented in Appendix B.

Research methodology selected for hypothesis testing. Out of varieties of MCDM tools, the present study selected a newly developed technique called BWM. As compared to other techniques, BWM is simpler, needs lesser data and yields more consistent and structured comparisons giving more reliable results according to Salimi and Rezaei (2018) who cited varieties of practical situations (e.g., risk assessment, supplier segmentation, and selection, supply chain management, water management, innovation management, etc.) where BWM was employed in the past. Recently, in their study on BT adoption in CE, Rejeb et al. (2022) also employed the BWM.for data analysis. In the present study, the EPASH problem was expressed as an MCDM problem as shown in Table 1. The objective is to select B*, the best (e.g., most valuable, most important) alternative out of the three perceptions of the considered three stakeholders of the auto industry. BWM uses pair-wise comparison to find the weights (w_j*) of the criteria and shows how much the decision-maker prefers criterion i over criterion j. The procedural steps followed were the same as Rezaei (2018), briefed as follows: (1) A set of decision criteria is determined. (2) The B(Best) and the W(Worst) criterion from experts' viewpoints are found (3) On the Likert scale (1 to 5), the preference of B over all others yields BO (best to others) vector. (4) Similarly, OW (others to worst) vector is determined. (5) Determine the optimal weight (w1*, ...,wn*). Thus, a Linear Programming problem, as briefed by Rejeb et al. (2022), is obtained and its solution yields the values of the optimal weights and the objective function providing the value of the consistency ratio (CR). The lower CR values imply more consistent and more reliable results. For MCDM problems, we can multiply the weights of different levels to determine the global weights.

Results and discussion

Weights of the yardsticks for the stakeholders' perceptions

With five IT experts, data were collected, BWM was applied and weights for experts were found to determine aggregate and overall weights for the criteria and sub-criteria as shown in Table 2.

Table 2

It was found that BTP and BTA impacts on CE are the most important yardsticks. In light of the local weightings, the most important, next most important, and least important sub-criteria were, respectively, found to be UBTB4 (uncertainty in CE-BT as per the users; weight: 0.263), UBTA5 (use of

Table 2. Weights for main criteria, sub-criteria, users (U), manufacturers (M), and dealers (D) and averages (Avg).

Main Criteria	Wei-ght (Wt)	Sub Criteria	Local Wt (LW) U	Global Wt (GW) U	LW M	GW M	LW D	GW D	Avg U	Avg M	Avg D
Potentials	0.393	BTP1	0.20	0.06	0.18	0.01	0.07	0.01	3.57	3.8	4
		BTP2	0.21	0.08	0.21	0.02	0.07	0.02	3.8	4.4	4
		BTP3	0.21	0.07	0.21	0.01	0.08	0.01	3.8	4.4	4.4
		BTP4	0.19	0.08	0.20	0.02	0.08	0.02	3.3	4.2	4.6
		BTP5	0.19	0.10	0.21	0.02	0.08	0.02	3.6	4.4	4.4
Barriers	0.262	BTB1	0.19	0.06	0.18	0.01	0.06	0.01	2.9	2.7	3.4
		BTB2	0.19	0.05	0.21	0.01	0.05	0.01	2.9	2.8	2.8
		BTB3	0.20	0.04	0.21	0.01	0.07	0.01	2.93	4	4
		BTB4	0.26	0.04	0.20	0.01	0.08	0.01	3.97	4.3	4.4
		BTB5	0.16	0.07	0.21	0.02	0.03	0.01	2.27	4	2
Applications	0.345	BTA1	0.16	0.06	0.21	0.01	0.07	0.01	2.8	4.4	4.2
		BTA2	0.20	0.07	0.22	0.02	0.08	0,015	3.53	4.6	44
		BTA3	0.21	0.09	0.21	0.02	0.07	0.02	3.73	4.4	4.2
		BTA4	0.20	0.07	0.20	0.01	0.07	0.01	3.43	4.2	4.2
		BTA5	0.23	0.06	0.18	0.01	0.06	0.01	4	3.8	3.6

BT in bitcoins; weight: 0.226) and DBTB5 (CE-BT requiring energy consumption and resource depletion as per the dealers; weight: 0.034). Graphically, these facts and other such features are illustrated in Figure 2.

In terms of the global weights, the most significant, the next most significant, and the least important sub-criteria were, respectively, found to be UBTP5 (CE-BT requiring energy consumption & resource depletion as per the users; weight: 0.100), UBTA3 (CE-BT can support all the UNO SDGs as per the users; weight: 0.093), MBTB2 (high cost of CE-BT impact as per the auto manufacturers) and DBTB5 (CE-BT impact requiring energy consumption & resource depletion as per the dealers) both carrying the same weight of 0.008. These findings reveal two important facts. First, the local weights and global weights do not match. Second, different stakeholders have different perceptions about the importance of the same sub-criteria. For example, users' category perceived UBTP5 to be most important whereas dealers perceived the same to be least important and the reasons for this deviation are obvious from users' and dealers' business viewpoints. When analyzed in terms of different stakeholders in the auto industry, the end users revealed that the sub-criteria, 'Application of BT in bitcoins,' is the most important one followed by the next one which is 'BT involves a lot of uncertainties. On the other hand, the sub-criteria, 'Larger digitization in BT requiring more energy consumption and resource depletion' emerged to be the least important perception among the users' category. Similarly, among the manufacturers' category, the sub-criteria 'BT used in manufacturing' was found to be of maximum importance while that of 'BT understanding in the industry' was found

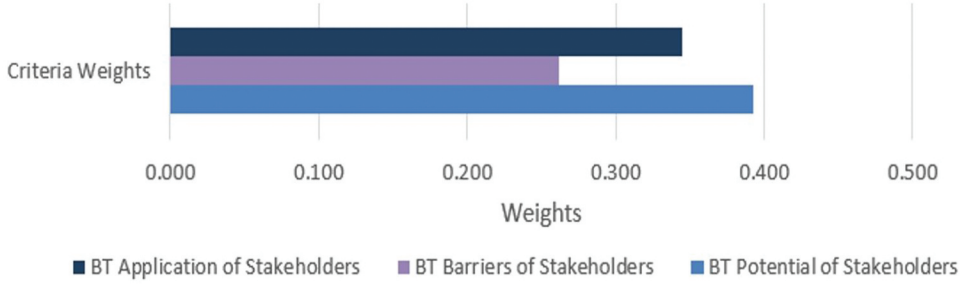


Figure 2. The relative importance of the primary or main criteria.

to be the least important. Among the automobile dealers, 'CE-BT impact enables ELV recovery' was found to carry the highest value while the sub-criteria 'BT requiring more energy consumption and resource depletion' was found to be the least important.

Stakeholders' CE-BT impact perceptions data analysis

The measurements of the perception scores and the weights of the yardsticks regarding CE-BT impact evaluation are given in Table 2. The aggregated BT impact evaluation scores of the various stakeholders with respect to their respective perceptions were computed as shown in columns 2, 4, and 6 of Table 4, while the ranking of each stakeholder's perception is presented in columns 3, 5, and 7. Furthermore, the overall aggregated BT evaluation of each stakeholder based on items of all perceptions and overall ranking based on this aggregated number are shown in columns 8 and 9 respectively. Assigning weights to different items (sub-criteria) and to different perceptions (main criteria) produces significant differences in the overall CE-BT impact evaluation scores for different stakeholders (Table 3). Finally, Table 4 depicts the CE-BT impact on three

stakeholders in totality. It was found that all the 45 stakeholders, considered, gave maximum weight to the 'Reuse dimension of CE' while the second one is the 'CE applications' that can be best tackled by using IT tools like BT. The third and fourth columns specify the three stakeholders' perspectives individually on the CE-BT impact and all gave maximum weight to the 'Return dimension of CE.' End users gave maximum weight to this dimension when compared to the other two (6.30). Also, all the three stakeholders assess the impact of CE-BT impacts' category 'CE Applications' as the second most important while the end-users of automobiles rated this dimension as the highest when compared to the other two. These facts are illustrated through the radar diagram presented in Figure 3.

Table 3. The overall ranking of the auto stakeholders' perception scores concerning the different criteria pertaining to BT use in the Indian auto industry. (U: Users; M: Manufacturers; D: Dealers).

Sub Crit	U (Score)	U (Rank)	M (Score)	M (Rank)	D (Score)	D (Rank)	Agr. All Score	Rank All
BTP1	3.57	6	3.80	5	4.00	4	3.79	10
BTP2	3.80	3	4.40	2	4.00	4	4.07	6
BTP3	3.80	3	4.40	2	4.40	2	4.20	2
BTP4	3.30	9	4.20	3	4.60	1	4.03	7
BTP5	3.60	5	4.40	2	4.40	2	4.13	4
BTB1	2.90	11	2.70	7	3.40	6	3.00	12
BTB2	2.90	11	2.80	6	2.80	7	2.83	13
BTB3	2.93	10	4.00	4	4.00	4	3.64	11
BTB4	3.97	2	4.30	3	4.40	2	4.22	1
BTB5	2.27	13	4.00	4	2.00	8	2.76	14
BTA1	2.80	12	4.40	2	4.20	3	3.80	9
BTA2	3.53	7	4.60	1	4.40	2	4.18	3
BTA3	3.73	4	4.40	2	4.20	3	4.11	5
BTA4	3.43	8	4.20	3	4.20	3	3.94	8
BTA5	4.00	1	3.80	5	3.60	5	3.80	9

Table 4. CE-BT nexus impact evaluation of the perception of manufacturers (M), dealers(D), and users (U) in the auto industry.

CE-BT Impact Features	Questionnaire Item Nos.	M's Persp.	D's Persp.	U's Persp.	Overall Average
CE Applications	1, 11, 12, 13, 14, 15	4.2	4.1	3.51	3.94
CE Limitations	5, 6, 8, 9	3.85	4.05	3.35	3.75
CE Cost	7, 10	3.4	2.4	2.58	2.79
Return Dimension of CE	2, 4	4.3	4.3	3.55	4.05
Reuse Dimension of CE	3	4.4	4.4	3.80	4.20

Persp. Refers to 'Perspectives' in the table

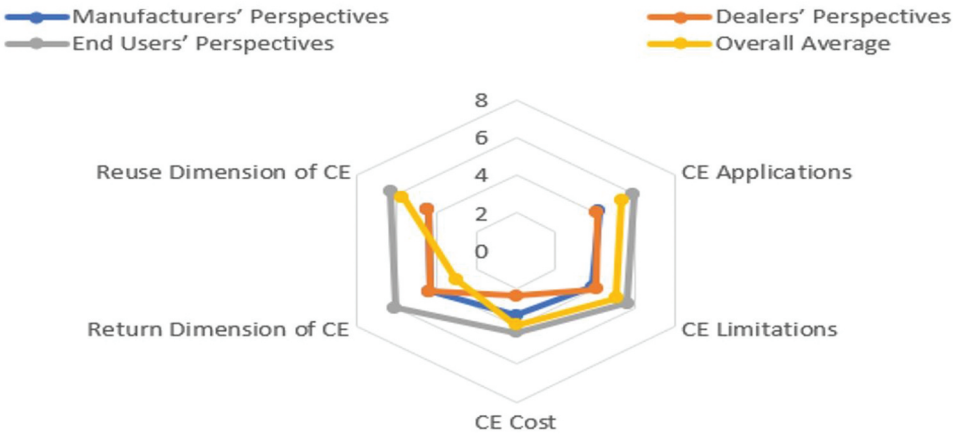


Figure 3. Radar diagram for the results of the study pertaining to CE-BT nexus impacts.

Discussion

The integration of circular economy and Industry 4.0 components like internet-of-things (IoT), blockchain technology (BT), etc. results in a faster induction of 'circularity' and leads to new strategies on natural resource utilization. Thus, the industrial world of tomorrow is likely to become more sustainable. It is reported in the literature that BT is a 'key enabler' in encountering the upcoming challenges in the CE applications and in the competitive industrial world of tomorrow, BT adoption in CE would be getting more critical. From the users of the newly inducted technologies and the corresponding stakeholders' viewpoints, evaluation of the impacts of CE-BT kind of interfaces plays a critical role in the technological developments. Also, the perception of 'technology acceptance by its users' concept is getting a highly significant status in the literature that revealed that there have been many studies in the past to study the role of human users 'perception and their perspectives in the development and acceptance of innovation and newer technologies from a socio-economic point of view. Recently, it was observed that 'perception-based' more empirical studies on technology acceptance were undertaken. For example, the Davis model called TAM (Technology Acceptance Model) remained well accepted and popular in its application, as opined by Chatterjee et al. (2021) who also employed this model in their study on the acceptance of AI technology adoption in the manufacturing environment. Similarly, Jayashankar et al. (2018) conducted the farmers' perception-based technology-adoption kind of study to explore how the trust in technology impacts the perceived risk in technology adoption. Along these lines, it was observed that the success of the application of any innovative technological effort is also a function of the perception of an organization's employees regarding their ability in performing the job associated with the incoming technologies like those of I4.0 On the other hand, the CE literature revealed the virtual absence of any such investigation in the past where the new emerging technology, BT, was explored from the stakeholders' perception of technology-use point of view. Thus, the present study might be the first of its kind in developing an understanding of the CE-BT nexus impacts on the Indian auto industry. Kamble et al. (2021) observed that BT positively affects sustainable supply chain performance, with supply chain integration as the intermediary variable between the two. Along similar lines, Khan et al. (2021) articulate that CE strategies such as circular procurement, circular design, recycling, and remanufacturing improved significantly through the application of BT on a large scale, particularly in the case of developing economies. However, such insights in the Indian context and that too on the Indian auto sector per se are by and large missing from

the literature. However, such or similar explorations undertaken in a non-Indian context, on non-BT digital tools and not related, particularly to the CE environment, are available in the literature. For example, it was observed that newer technologies like AI might be easier to adopt when support from the top level of management is obtained (Pu et al. 2019). Present findings of the manufacturers' perception regarding CE-BT impacts become all the more important when it is noted that, first, the impact of the applications of AI, is not well understood in the field of manufacturing (Chen et al. 2020). Secondly, in terms of the potentialities of AI on the enterprise level, not many investigations were taken in the past (Pu et al. 2019). It might be noted that compared to AI, BT is a more recently evolved digital tool, yet AI appears to be so less explored a field. Thus, studying CE-BT impact on the auto represents relatively a new source of knowledge enrichment in the area of CE-BT applications in the presently available literature on the theme. de Graaf, Allouch, and van Dijk (2017) found that the 'usefulness' and 'intention' were linearly related to each other. Thus, it appears that the term 'perception' might be taken as a function of two basic factors: technology acceptance and users' intentions in the context of the 'technology use' as suggested by Cagliano et al. (2019). As discussed earlier, the TAM model appears to be catering well and could explain a minimum of 40% of the causes responsible for the technology utilization by the users. In the present study, both the 'Reuse' strategy and 'CE-BT applications' were found to be among the most crucial CE-BT impacts. Similar findings have been articulated by Singh and Gurtu (2021) in the manufacturing context. In support of this prerogative, it can further be added that when one of the nodes in the centralized information systems fails, BT mitigates this huge risk by allowing the smart manufacturing systems to operate normally (Song and Zhu 2021). Also, BT can contribute to CE through its foolproof features like decentralization, distributive, and high security. On the other hand, it is felt that 'cutting edge technologies' are offered by the I4.0 era and BT being one of the constituents of I4.0, the CE-BT nexus impacts need to be explored in the days ahead. In this light, the present study might be taken as a good initiative in the auto industry for its future benefits by way of the CE-BT applications in the field. At the same, time there are some limitations and barriers also which have, of late, appeared to be emerging in literature. For example, in the path of BT adoption in CE, Rejeb et al. (2022) identified some major barriers as 'lack of knowledge and management support, reluctance to change, and technological immaturity' at the global level with some of them, also supported by Awan and Sroufe (2022), especially in the US case.

Conclusion and directions for future research

In the ever-expanding era of digitization in the industrial world, all the manufacturing functions including automaking are getting revolutionized day by day. Under the impact of the rapid growth of industrial production, material scarcity is also going up exponentially, thereby demanding the immediate need for the 'circular economy.' In light of such a scenario, the present study was undertaken as discussed above. Based on it, the following conclusions are drawn:

- The study proposes a new CE-BT model that appears to be more reliable and secure than other classical models applied to CE in the auto sector.
- The study revealed that all the three stakeholders considered, perceived 'Reuse dimension' and 'CE-BT applications' among the most crucial CE-BT impacts on the Indian auto industry. Also, different criteria that were considered weighted differently as perceived by the stakeholders.
- Data pertaining to CE revealed the 'potential' and 'applications' as the most critical.
- This knowledge might be utilized by the future planners of the auto industry in knowing the intentions of the I4.0 like 'technology acceptance' of the respective stakeholders.
- This might be the first attempt of the kind in the CE-BT literature and it may help the future planners of the digitized CE induction in the Indian auto industry.

Future researchers might address the issues of:

- comparative perception of different stakeholders of the auto industry.
- similar studies for non-auto sectors, non-BT I4.0 tools, and other non-Indian parts of the world as well.
- In future studies, the perception data might be more enriched by including a larger number of auto stakeholders. Similarly, the list of measures of CE-BT impact factors might include more numbers in future studies.
- Also, the modeling may further be improved in future studies by way of employing such techniques as System Dynamics Modeling (SDM).

Highlights

- The paper explores the perception of the Indian auto stakeholders on the impact of blockchain technology in the circular economy context.
- A new circular economy-blockchain model and a case study based on BWM are presented.
- Data analysis revealed 'potential' as the most and 'barriers' as the least important impacts.
- Proposed model appears to be reliable and secure.
- Findings might help the future planners of blockchain technology in the auto industry.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Syed Wasiul Hasan Rizvi is currently a PhD scholar at Delhi Technological University (DTU) in the Department of Mechanical Engineering. E-Mail id: wasiulhasan@yahoo.com

Saurabh Agrawal is Associate Professor in Delhi School of Management, DTU, Delhi, India. E-Mail ID: agrawals.iit@gmail.com

Qasim Murtaza is a Professor in the Department of Mechanical Engineering, DTU, Delhi, India. E-Mail ID: qasimmurtaza@gmail.com

ORCID

Syed Wasiul Hasan Rizvi http://orcid.org/0000-0002-4653-4337
Qasim Murtaza http://orcid.org/0000-0001-8784-2304

References

Alhawari, O., U. Awan, M. K. S. Bhutta, M. A. Ülkü, et al. 2021. Insights from circular economy literature: A review of extant definitions and unravelling paths to future research. *Sustainability* 13 (2859):859. doi:10.3390/su13020859.

Awan, U., R. Sroufe, and M. Shahbaz. 2021. Industry 4.0 and the circular economy: A literature review and recommendations for future research. *Business Strategy and Environment* 30 (4 2038–2060). doi:10.1002/bse.2731.

Awan, U., Bhatti, S. H., Shamim, S., Khan, Z., Akhtar, P., and Balta, M. E., et al. 2021a. The role of big data analytics in manufacturing agility and performance: Moderation–mediation analysis of organizational creativity and of the involvement of customers as data analysts. *British Journal of Management* 33(3): 1200–1220. doi:10.1111/1467-8551.12549.

Awan, U., Shamim, S., Khan, Z., Zia, N. U., Shariq, S. M., and Khan, M. N., et al. 2021b. Big data analytics capability and decision-making:The role of data-driven insight on circular economy performance, technological forecasting 168:120766. doi:10.1016/j.techfore.2021.120766.

Awan, U., and R. Sroufe. 2022. Sustainability in the circular economy: Insights and dynamics of designing circular business models. *Applied Sciences* 12 (3):1521. doi:10.3390/app12031521.

Bumblauskas, D., A. Mann, B. Dugan, J. Rittmer. 2019. A blockchain use case in food distribution: Do you know where your food has been? *International Journal of Information Management* 52:102008. doi:10.1016/j.ijinfomgt.2019.09.004.

Cagliano, R., F. Canterino, A. Longoni, and E. Bartezzaghi. 2019. The interplay between smart manufacturing technologies and work organization: The role of technological complexity. *International Journal of Operations & Production Management* 39 (6/7/8):913–934. doi:10.1108/IJOPM-01-2019-0093.

Chatterjee, S., N. P. Rana, Y. K. Dwivedi, A. M. Baabdullah. 2021. Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model. *Technological Forecasting and Social Change* 170:120880. doi:10.1016/j.techfore.2021.120880.

Chen, J., M. Gao, S. K. Mangla, M. Song, J. Wen. 2020. Effects of technological changes on China's carbon emissions. *Technological Forecasting and Social Change* 153:119938. doi:10.1016/j.techfore.2020.119938.

Cheng, Y., U. Awan, S. Ahmad, Z. Tan. 2021. How do technological innovation and fiscal decentralization affect the environment? A story of the fourth industrial revolution and sustainable growth. *Technological Forecasting and Social Change* 162:120398. doi:10.1016/j.techfore.2020.120398.

Dalenogare, L., G. B. Benitez, N. F. Ayala, and A. G. Frank. 2018. The expected contribution of Industry 4.0 technologies for industrial performance. *International Journal of Production Economics* 204:383–394. doi:10.1016/j.ijpe.2018.08.019.

de Graaf, M. M. A., S. B. Allouch, and J. A. G. M. van Dijk. 2017. A phased framework for long-term user acceptance of interactive technology in domestic environments. *New Media Society* 20 (2):146144481772726. doi:10.1177/1461444817727264.

Gerger. 2021. Use cases and statistical evaluation, in blockchain technology in the automotive industry. *First Ed. IGI Global*. doi:10.4018/978-1-7998-6650-3.ch012.

Gigauri. 2021. New economic concepts shaping business models in post-pandemic era. *International Journal of Innovative Technologies in Economy*. doi:10.31435/rsglobal_ijite/30032021/7393.

Govindan, K., and M. Hasanagic. 2018. A systematic review on drivers, barriers, and practices towards circular economy: A supply chain perspective. *International Journal of Production Research* 56 (1–2):278–311. doi:10.1080/00207543.2017.1402141.

Jayashankar, P., S. Nilakanta, W. J. Johnston, P. Gill, and R. Burres. 2018. IoT adoption in agriculture: The role of trust, perceived value and risk. *The Journal of Business and Industrial Marketing* 33 (6):804–821. doi:10.1108/JBIM-01-2018-0023.

Kamble, S. S., A. Gunasekaran, N. Subramanian, A. Ghadge, A. Belhadi, M. Venkatesh. 2021. Blockchain technology's impact on supply chain integration and sustainable supply chain performance: Evidence from the automotive industry. *Annals of Operations Research*. doi:10.1007/s10479-021-04129-6.

Khan, S. A. R., A. Razzaq, Z. Yu, and S. Miller. 2021. Industry 4.0 and circular economy practices: A new era business strategies for environmental sustainability. *Business Strategy and the Environment* 30 (8):4001–4014. doi:10.1002/bse.2853.

Kirchherr, J., D. Reike, and M. Hekkert. 2017. Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling* 127:221–232. doi:10.1016/j.resconrec.2017.09.005.

Kshetri, N. 2018. Blockchain's roles in meeting key supply chain management objectives. *International Journal for Information Management* 39:80–89. doi:10.1016/j.ijinfomgt.2017.12.005.

MacArthur, E. D., and D. Waughray, 2016. Accessed Feb 10 2016. Intelligent assets: Unlocking the circular economy potential. https://www.ellenmacarthurfoundation.org/publications/intelligent-assets

Okorie, O., K. Salonitis, F. Charnley, M. Moreno, C. Turner, and A. Tiwari. 2018. Digitisation and the circular economy: A review of current research and future trends. *Energies* 11 (11):3009. doi:10.3390/en11113009.

Pu, X., A. Y. L. Chong, Z. Cai, M. K. Lim, and K. H. Tan. 2019. Leveraging open-standard interorganizational information systems for process adaptability and alignment: An empirical analysis. *International Journal of Operations and Production Management* 39 (6/7/8):962–992. doi:10.1108/IJOPM-12-2018-0747.

Rashid Khan, H., U. Awan, K. Zaman, A. A. Nassani, M. Haffar, M. M. Q. Abro. 2021. Assessing hybrid solar-wind potential for industrial decarbonization strategies: Global shift to green development. *Energies* 14 (22):7620. doi:10.3390/en14227620.

Rayome. 2019. Top 10 emerging technologies of 2019. *TechRepublic*. Accessed Apr 25 2022. https://www.techrepublic.com/article/top-10-emerging-technologies-of-2019.

Rejeb, A. 2022. Barriers to blockchain adoption in the circular economy: A Fuzzy Delphi and best-worst approach. *Sustainability* 143611. doi:10.3390/su14063611.

Rizvi, S. W. H., S. Agrawal, and Q. Murtaza. 2021. Circular economy under the impact of IT tools: A content-based review. *International Journal of Sustainability Engineering* 14 (2):87–97. doi:10.1080/19397038.2020.1773567.

Salimi, N., and J. Rezaei. 2018. Evaluating firms' R&D performance using best worst method. *Evaluation and Program Planning* 66:147–55. doi:10.1016/j.evalprogplan.2017.10.002.

Schinckus, C. 2020. The good, the bad and the ugly: An overview of the sustainability of blockchain technology. *Energy Research & Social Science* 69 (101614). doi:10.1016/j.erss.2020.101614.

Singh, R. K., and A. Gurtu. 2021. Embracing advanced manufacturing technologies for performance improvement: An empirical study. *Benchmarking: An International Journal*. doi:10.1108/BIJ-02-2021-0110.

Song, Z., and J. Zhu. 2021. Blockchain for smart manufacturing systems: A survey. *Chinese Management Studies*. doi:10.1108/CMS-04-2021-0152.

Upadhyayet al , 2021. Blockchain Technology and the Circular Economy: Implications for Sustainability and Social Responsibility, Journal of Cleaner Production 293: 126130. doi: 10.1016/j.jclepro.2021.126130

Wang, S., Ouyang, L., Yuan, Y., Ni, X., Han, X., and Wang, F., et al. 2019. Blockchain-enabled smart contracts: Architecture, applications, and future trends. *IEEE Transactions on Systems Man and Cybernetics: Systems* 49 (11):2266–2277. doi:10.1109/TSMC.2019.2895123.

World Economic Forum White Paper, 2016. Accessed Apr 24 2022. Digital transformation of industries: Digital enterprise. https://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/digital-enterprise-narrative-final-january-2016.pdf

Yadav, V. S., A. R. Singh, R. D. Raut, U. H. Govindarajan. 2020. Blockchain technology adoption barriers in the Indian agricultural supply chain: An integrated approach. *Resources, Conservation and Recycling* 161:104877. doi:10.1016/j.resconrec.2020.104877.

Zhang, D. 2019. Application of blockchain technology in incentivizing efficient use of rural wastes: A case study on yitong system. *Energy Procedia* 158:6707–6714. doi:10.1016/j.egypro.2019.01.018.
Zhang, Y., and U. Guin. 2020. End-to-End traceability of ICs in component supply chain for fighting against recycling. *IEEE Transactions Information Forensics and Security* 15:767–75. doi:10.1109/TIFS.2019.2928493.
Zhao, G., S. Liu, C. Lopez, H. Lu, S. Elgueta, H. Chen, B. M. Boshkoska. 2019. Blockchain technology in agri-food value chain management: A synthesis of applications, challenges and future research directions. *Computers in Industry* 109:83–99. doi:10.1016/j.compind.2019.04.002.

Appendix-A: The Research Instrument (Questionnaire)

1. CE-BT nexus is drawing considerable attention from academia and industry.
2. BT enables effective 'traceability' of parts/products on the shop floor.
3. BT supports the 'reuse' strategy of CE/Sustainability
4. BT enables ELV (End of Life Vehicles) recovery.
5. The data storage and interaction functionalities needed for CE are seldom employed.
6. Industry persons may not understand the CE-BT nexus.
7. Cost of BT applications in CE is high.
8. BT being a new digital tool, its real impact on CE cannot be assessed.
9. Studying BT impacts on CE involves a lot of uncertainties.
10. Larger digitization like that in BT needs more energy consumption and resource depletion.
11. BT is a digital technology that mainly refers to modern information and communication technology tools that might significantly affect the CE framework.
12. CE-BT has been reported to have been applied in the auto industry. (CE Applications)
13. BT can support all the SDGs (Sustainability Development Goals) set by UNO.
14. CE-BT nexus can support managerial decision making in the auto industry.
15. BT which has been reported to have been initially applied in the 'Bitcoin currency' field is expected to be impacting industry also on a large scale.

Appendix-B: Respondents' profiles

(i) For manufacturers of vehicles/parts

Turnover (in Rupees)	#	*No. of Employees*	#	*India Region*	No. of firms
Less than 100 Crores	0	Less than 50	0	North	7
100-500 Crores	4	50-200	0	East	0
More than 500 Crores	6	200-500	0	West	0
		More than 500	10	South	3

(ii) For end-users of vehicles

Employers' Turnover	No. of firms
Less than Rs.100 Crores	9
Rs. 100-500 Crores	10

(iii) For dealers/suppliers of vehicles

Dealers' Turnover	No. of firms
Less than Rs. 100 Crores	3
Rs. 100-500 Crores	1

arXiv:2208.06638v1 [math.CV] 13 Aug 2022

COEFFICIENT PROBLEMS FOR CERTAIN CLOSE-TO-CONVEX FUNCTIONS

MRIDULA MUNDALIA AND S. SIVAPRASAD KUMAR

Abstract. In this paper, sharp bounds are established for the second Hankel determinant of logarithmic coefficients for normalised analytic functions satisfying certain differential inequality.

1. Introduction

Let $\mathcal{A}$ be the class of all analytic functions defined on the open unit disc $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ with the Taylor series expansion of the form $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$. Assume $\mathcal{S} \subset \mathcal{A}$ to be the class of univalent functions defined on $\mathbb{D}$. A function $f \in \mathcal{S}$, lies in $\mathcal{S}^*$ if the domain $f(\mathbb{D})$ is starlike w.r.t origin. A function $f \in \mathcal{A}$ belongs to the class of close-to-convex functions $\mathcal{K}$ [14], if there exists $g \in \mathcal{S}^*$, such that $\mathrm{Re}(e^{i\beta} z f'(z)/g(z)) > 0$ for $z \in \mathbb{D}$ and $\beta \in (-\pi/2, \pi/2)$. Note that $\mathcal{S}^* \subset \mathcal{K} \subset \mathcal{S}$. Moreover, for $\beta = 0$ and specific choices of $g(z)$, namely $g(z) = 1/(1-z), 1/(1-z^2), 1/(1-z+z^2)$ and $1/(1-z)^2$, we obtain some special subclasses of close-to-convex functions $\mathcal{F}_i$ $(i = 1, \ldots, 4)$, defined as

$$\begin{aligned} \mathcal{F}_1 &:= \left\{ f \in \mathcal{A} : \mathrm{Re}(1-z) f'(z) > 0 \right\} \\ \mathcal{F}_2 &:= \left\{ f \in \mathcal{A} : \mathrm{Re}(1-z^2) f'(z) > 0 \right\} \\ \mathcal{F}_3 &:= \left\{ f \in \mathcal{A} : \mathrm{Re}(1-z+z^2) f'(z) > 0 \right\} \\ \mathcal{F}_4 &:= \left\{ f \in \mathcal{A} : \mathrm{Re}(1-z)^2 f'(z) > 0 \right\}, \end{aligned} \tag{1.1}$$

where $z \in \mathbb{D}$. A function $f \in \mathcal{S}_s^*$ if for any $r < 1$ sufficiently close to 1, and any γ lying on the circle $|z| = r$, the angular velocity of $f(z)$ about the point $f(-\gamma)$ is positive at γ as z traverses the circle $|z| = r$ in the positive direction, i.e $\mathrm{Re}(2zf'(z)/(f(z) - f(-\gamma))) > 0$ for $|z| = r$ at $z = \gamma$. In 1959, Sakaguchi [30] introduced and examined the class $\mathcal{S}_s^*$, consisting of functions starlike with respect to symmetric points, characterised by

$$\mathrm{Re}\left(\frac{2zf'(z)}{f(z) - f(-z)} \right) > 0 \quad (z \in \mathbb{D}). \tag{1.2}$$

Note that Sakaguchi [30] proved that $\mathcal{S}_s^* \subset \mathcal{K}$. The bounds of γ_n for functions in $\mathcal{K}$ were examined in [32, p. 116], [31]. In 2018, Kumar and Vasudevrao [17] obtained bounds on early logarithmic coefficients for the subclasses $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ of $\mathcal{K}$. Infact in [1] the estimates of logarithmic coefficients of class $\mathcal{F}_4 \subset \mathcal{S}$ was discussed. Note that the conditions mentioned in classes $\mathcal{F}_1, \mathcal{F}_2$ and $\mathcal{F}_4$ were considered by Ozaki [27] as an important criteria of univalence. Moreover, it is important to mention that, the classes $\mathcal{F}_2$ and $\mathcal{F}_4$ have nice geometrical interpretations. Each member of $\mathcal{F}_2$ is convex in the direction of imaginary axis and every $f \in \mathcal{F}_4$ is convex in the positive direction of real axis. For $f \in \mathcal{S}$, define $\widetilde{F}_f(z) := 2F_f(z)$, where

$$F_f(z) = \log\left(\frac{f(z)}{z} \right) = 2\sum_{n=1}^{\infty} \gamma_n(f) z^n \quad (z \in \mathbb{D}\backslash\{0\}, \log 1 := 0). \tag{1.3}$$

2010 *Mathematics Subject Classification.* 30C45, 30C80.

Key words and phrases. Univalent functions, Starlike functions, Radius problems, Differential subordination.

The first author is supported by Delhi Technological University, New Delhi.

The coefficients $\gamma_n := \gamma_n(f)$ in (1.3) are known as logarithmic coefficients of f. Sharp logarithmic coefficient estimates for the class $\mathcal{S}$, are already known for $n = 1$ and $n = 2$, given by $|\gamma_1| \leq 1$ and $|\gamma_2| \leq 1/2 + 1/e^2$. However bound of γ_n for $n \geq 3$, is still an open problem. Logarithmic coefficients played a crucial role in Milin's conjecture[24], which states that if $f \in \mathcal{S}$, then

$$\sum_{m=1}^{n} \sum_{k=1}^{n} \left(k|\gamma_k|^2 - \frac{1}{k} \right) \leq 0.$$

For $q, n \in \mathbb{N}$, the q^{th} Hankel determinant $H_{q,n}(f)$, for $f \in \mathcal{A}$ is defined as,

$$H_{q,n}(f) = \begin{vmatrix} a_n & a_{n+1} & \cdots & a_{n+q-1} \\ a_{n+1} & a_{n+2} & \cdots & a_{n+q} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n+q-1} & a_{n+q} & \cdots & a_{n+2(q-1)} \end{vmatrix}.$$

Hankel determinant of exponential polynomials was studied by Ehrenborg in [8]. Hayman [20] studied some properties of Hankel transform of an integer sequence. Hankel determinants $H_{q,n}(f)$ are useful, in showing that a function of bounded characteristic in $\mathbb{D}$, i.e., a function which is a ratio of two bounded analytic functions with its Laurent series around the origin having integral coefficients, is rational [2]. It is worth mentioning that the Hankel determinant $H_{q,n}(f)$ has been investigated by several authors to study its rate of growth as $n \to \infty$. Additionally, the notion of Hankel determinant is extremely helpful in the study of power series with integral coefficients and in the theory of singularities [7]. Noonan and Thomas [25] studied the growth rate of the second Hankel determinant of an areally mean p-valent function. Pommerenke [28] proved that the Hankel determinants of univalent functions satisfy $|H_{q,n}(f)| < Kn^{-(\frac{1}{2}+\beta)q+\frac{3}{2}}$, where $\beta > 1/4000$ and K depends only on q. Later, Hayman [12] proved that $H_{2,n}(f) < An^{1/2}$, where A is an absolute constant, for areally mean univalent functions. One can observe that $H_{2,1}(f) = a_3 - a_2^2$, is the Fekete-Szegö functional, which is further generalised to $a_3 - \mu a_2^2$, where μ is a complex number. In 1916, Bieberbach [11] estimated $H_{2,1}(f)$ for the class $\mathcal{S}$. Several authors have studied second Hankel determinant (see [16, 21, 25]), denoted by $H_{2,2}(f)$ for certain subclasses of $\mathcal{A}$. Many authors have studied the problem of calculating $\max_{f \in \mathbb{F}} |H_{2,2}(f)|$ for various subfamilies $\mathbb{F} \subset \mathcal{A}$ [13, 15, 21]. Infact Noor [26] studied Hankel determinant for close-to-convex functions. For $f \in \mathcal{A}$, define the q^{th} Hankel determinant $H_{q,n}(F_f)$, where $q, n \in \mathbb{N}$, with entries as logarithmic coefficients,

$$H_{q,n}(F_f) = \begin{vmatrix} \gamma_n & \gamma_{n+1} & \cdots & \gamma_{n+q-1} \\ \gamma_{n+1} & \gamma_{n+2} & \cdots & \gamma_{n+q} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{n+q-1} & \gamma_{n+q} & \cdots & \gamma_{n+2(q-1)} \end{vmatrix}.$$

It is worth mentioning that, the problem of studying Hankel determinant with entries as logarithmic coefficients, was coined by Kowalczyk et al. [16]. In this paper, an attempt is made to study the problem of establishing sharp bound of $H_{2,1}(F_f) = \gamma_1\gamma_3 - \gamma_2^2$, which has a striking resemblance to $H_{2,1}(f) = a_2a_4 - a_3^2$, where f lies in some subclass of $\mathcal{A}$. Further, from (1.3) logarithmic coefficients of $f \in \mathcal{S}$, can be expressed as

$$\gamma_1 = \frac{1}{2}a_2, \quad \gamma_2 = \frac{1}{2}\left(a_3 - \frac{1}{2}a_2^2\right), \quad \gamma_3 = \frac{1}{4}\left(a_4 - a_2a_3 + \frac{1}{3}a_2^3\right).$$

Then for $f \in \mathcal{S}$

$$H_{2,1}(F_f) = \gamma_1\gamma_3 - \gamma_2^2 = \frac{1}{4}\left(a_2a_4 - a_3^2 + \frac{1}{12}a_2^4\right).$$

Note that for $f \in \mathcal{A}$, the functional $|H_{2,1}(F_f)| = |\gamma_1\gamma_3 - \gamma_2^2|$ is rotationally invariant. For this reason, it is sufficient to assume that the second coefficient a_2 is real. Several authors have studied

and established varied results for these classes (see [4, 5]). Based on the definition of each class in (1.1) and (1.2), functions in these classes can be represented in terms of Carathéodory class $\mathcal{P}$, consisting of analytic functions $p(z)$, defined as

$$p(z) = 1 + \sum_{n=1}^{\infty} c_n z^n \tag{1.4}$$

such that $\operatorname{Re} p(z) > 0$. For the purpose of computing the bounds for $H_{2,1}(F_f)$, we use the coefficient formula given in (1.5) for c_1 given in [3], for c_2 [29, pg. 166] and the formula for coefficient c_3 due to Libera and Zlotkiewicz [22, 23]. Moreover, results stated by Cho et al. [5] highlight the importance of functions given in (1.7) and (1.8). Classes defined in (1.1) and (1.2), have been extensively studied on different aspects earlier, namely early logarithmic and inverse logarithmic coefficients bounds [9], successive inverse coefficients [18], Hermitian-Toeplitz determinants [19]. However, no work has been carried out for Hankel determinants of logarithmic coefficients for functions lying in these classes. In this article, we will address this problem. Lemmas stated below serve as a prerequisite for deriving our subsequent results.

Lemma 1.1. [22, 23, 29] *If $p \in \mathcal{P}$ of the form $p(z) = 1 + c_1 z + c_2 z^2 + c_3 z^3 + \ldots$, with $c_1 \geq 0$, then*

$$\begin{aligned} c_1 &= 2\zeta_1 \\ c_2 &= 2\zeta_1^2 + 2(1-\zeta_1^2)\zeta_2 \end{aligned} \tag{1.5}$$

and

$$c_3 = 2\zeta_1^3 + 4(1-\zeta_1^2)\zeta_1\zeta_2 - 2(1-\zeta_1^2)\zeta_1\zeta_2^2 + 2(1-\zeta_1^2)(1-|\zeta_2|^2)\zeta_3 \tag{1.6}$$

for some $\zeta_1 \in [0,1]$ and $\zeta_2, \zeta_3 \in \overline{\mathbb{D}} := \{z \in \mathbb{C} : |z| < 1\}$. For $\zeta_1 \in \mathbb{D}$ and $\zeta_2 \in \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, there is a unique function $p \in \mathcal{P}$ with c_1 and c_2 as in (1.5)-(1.6), namely

$$p(z) = \frac{1 + (\overline{\zeta_1}\zeta_2 + \zeta_1)z + \zeta_2 z^2}{1 + (\overline{\zeta_1}\zeta_2 - \zeta_1)z - \zeta_2 z^2} \quad (z \in \mathbb{D}). \tag{1.7}$$

For $\zeta_1, \zeta_2 \in \mathbb{D}$ and $\zeta_3 \in \mathbb{T}$, there exists a unique $p \in \mathcal{P}$ with c_1, c_2 and c_3 as in (1.5)-(1.6), namely,

$$\frac{1 + (\overline{\zeta_1}\zeta_2 + \overline{\zeta_2}\zeta_3 + \zeta_1)z + (\overline{\zeta_1}\zeta_3 + \zeta_1\overline{\zeta_2}\zeta_3 + \zeta_2)z^2 + \zeta_3 z^3}{1 + (\overline{\zeta_1}\zeta_2 + \overline{\zeta_2}\zeta_3 - \zeta_1)z + (\overline{\zeta_1}\zeta_3 - \zeta_1\overline{\zeta_2}\zeta_3 - \zeta_2)z^2 - \zeta_3 z^3} \quad (z \in \mathbb{D}). \tag{1.8}$$

Following result is due to Choi, Kim and Sugawa [6].

Lemma 1.2. [6] *For $A, B, C \in \mathbb{R}$, let*

$$Y(A,B,C) := \max\left\{|A + Bz + Cz^2| + 1 - |z|^2 : z \in \overline{\mathbb{D}}\right\}.$$

(1) If $AC \geq 0$, then

$$Y(A,B,C) = \begin{cases} |A| + |B| + |C|, & |B| \geq 2(1-|C|), \\ 1 + |A| + \dfrac{B^2}{4(1-|C|)}, & |B| < 2(1-|C|). \end{cases}$$

(2) If $AC < 0$, then

$$Y(A,B,C) = \begin{cases} 1-|A|+\dfrac{|B|}{4(1-|C|)}, & -4AC(C^{-2}-1) \leq B^2 \wedge |B| < 2(1-|C|), \\ 1+|A|+\dfrac{B^2}{4(1+|C|)}, & B^2 < \min\{4(1+|C|^2), -4AC(C^{-2}-1)\}, \\ R(A,B,C), & \textit{otherwise} \end{cases}$$

where

$$R(A,B,C) = \begin{cases} |A|+|B|-|C|, & |C|(|B|+4|A|) \leq |AB|, \\ -|A|+|B|+|C|, & |AB| \leq |C|(|B|-4|A|), \\ (|C|+|A|)\sqrt{1-\dfrac{B^2}{4AC}}, & \textit{otherwise.} \end{cases}$$

2. Hankel Determinant of Logarithmic Coefficients

In this section, we establish results pertaining to sharp estimates of $H_{2,1}(F_f)$ for the classes $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4$ and $\mathcal{S}_s^*$. We begin by determining sharp bounds of $H_{2,1}(F_f)$ for $f \in \mathcal{S}_s^*$.

Theorem 2.1. *Let F_f be given by* (1.3)*. If $f \in \mathcal{S}_s^*$, then sharp bound on Hankel determinant for F_f is given by*

$$|H_{2,1}(F_f)| \leq \frac{1}{4}. \tag{2.1}$$

Above inequality is sharp due to the function

$$\tilde{f}(z) = \int_0^z (1+t^2)/(1-t^2)^2 dt. \tag{2.2}$$

Proof. Suppose $f \in \mathcal{A}$ of the form $f(z) = z + a_2 z^2 + a_3 z^3 + a_4 z^4 + \cdots$ satisfy

$$\frac{2zf'(z)}{f(z)-f(-z)} = p(z), \tag{2.3}$$

where $p \in \mathcal{P}$ is given by (1.4). It is a well-known fact that $\mathcal{P}$ is invariant under rotation, we assume $c_1 \in [0,2]$. From equation (2.3) we express coefficients of $f(z)$, $a_i's$ $(i = 2,3,4)$ in terms of $c_i's$ $(i = 1,2,3)$ as follows

$$a_2 = \frac{1}{2}c_1, \quad a_3 = \frac{1}{2}c_2, \quad a_4 = \frac{1}{8}\left(c_1c_2 - 2c_3\right).$$

Lemma 1.1 yields the following expression in terms of ζ_i $'s$, where $\zeta_i \in \overline{\mathbb{D}}$ $(i = 1,2,3)$.

$$\begin{aligned} \mathcal{L} := \gamma_1\gamma_3 - \gamma_2^2 &= \frac{1}{4}\left(a_4a_2 - a_3^2 + \frac{1}{12}a_2^4\right) \\ &= \frac{1}{768}(c_1^4 + 12c_2c_1^2 - 24c_3c_1 - 48c_2^2) \\ &= \frac{1}{48}(6\zeta_2^2(5\zeta_1^2 - 3\zeta_1^4 - 2) - 11\zeta_1^4 - 6\zeta_1(1-\zeta_1^2)\zeta_3(1-|\zeta_2|^2) \\ &\quad - 30\zeta_2\zeta_1^2(1-\zeta_1^2)). \end{aligned} \tag{2.4}$$

Expression in (2.4) leads to,

$$|\mathcal{L}|= \begin{cases} 1/4, & \zeta_1 = 0, \\ 11/48, & \zeta_1 = 1. \end{cases}$$

For $\zeta_1 \in (0,1)$ and inequality $|\zeta_3| \leq 1$, the expression in (2.4) together with Lemma 1.1 results in the following inequality

$$|\mathcal{L}| \leq \frac{1}{8}\zeta_1(1-\zeta_1^2)\Psi(A,B,C) \tag{2.5}$$

where

$$\Psi(A,B,C) := |A + B\zeta_2 + C\zeta_2^2| + 1 - |\zeta_2|^2$$

with

$$A = -\frac{11\zeta_1^3}{6\left(1-\zeta_1^2\right)}, \quad B = -5\zeta_1, \quad C = 3\zeta_1 - \frac{2}{\zeta_1}. \tag{2.6}$$

In view of Lemma 1.2, we now examine the following cases based on the expressions of A, B and C given in (2.6).

I. Let $\zeta_1 \in X = (0, \zeta^*]$, where $\zeta^* = \sqrt{2/3}$. It is easy to verify that $AC \geq 0$. Since $|B| - 2(1-|C|) = \left(4 - \zeta_1(2+\zeta_1)\right)/\zeta_1$ is an increasing function on X, then $|B| - 2(1-|C|) \geq 5\zeta^* - 2 > 0$. Thus on applying Lemma 1.1 to inequality (2.5), we have

$$|\mathcal{L}| \leq \frac{1}{8}\zeta_1(1-\zeta_1^2)(|A| + |B| + |C|) = \frac{1}{4} - \frac{\zeta_1^4}{48} \leq \frac{1}{4}.$$

II. Assume $X^* = X^c \cap (0,1)$. Observe that $AC < 0$ for each $\zeta_1 \in X^*$. This case demands the following subcases.

A. Simple calculation reveals that, for each ζ_1 in X^*,

$$\begin{cases} T_1(\zeta_1) := |B| - 2(1-|C|) = \dfrac{1}{\zeta_1}(11\zeta_1^2 - 2\zeta_1 - 4) > 0 \\ \\ T_2(\zeta_1) := -4AC\left(\dfrac{1}{C^2} - 1\right) - B^2 = \dfrac{27\zeta_1^2 - 62}{\zeta_1^2(6 - 9\zeta_1^2)} < 0. \end{cases}$$

From above we conclude that the set $T_1(X^*) \cap T_2(X^*)$ is empty. Thus according to Lemma 1.2, this case fails to exist for any $\zeta_1 \in X^*$.

B. For $\zeta_1 \in X^*$, the expressions $4(1+|C|)^2$ and $-4AC(C^{-2} - 1)$ become

$$4\ T_3(\zeta_1) := 4(1+|C|)^2 \text{ and } T_4(\zeta_1) := -4AC\left(\frac{1}{C^2} - 1\right) = \frac{22\zeta_1^2(9\zeta_1^2 - 4)}{3(3\zeta_1^2 - 2)},$$

respectively, where $T_3(x) := (3x^2 + x - 2)^2/x^2$. It is easy to observe that $T_3'(x)$ is non-vanishing on X^*. Infact $T_3'(1) = 20$, therefore $T_3(x) < 4$. Since $\zeta^* > \zeta^{*2}$, this yields that $T_4(x)$ is positive on X^*. Thus inequality below is false for any $\zeta_1 \in X^*$,

$$B^2 = 25\zeta_1^2 < \min\{4\ T_3(x), T_4(x)\} = 4\ T_3(x).$$

C. The inequality below holds for $\zeta_1 \in X^*$,

$$|C|(|B| + 4|A|) - |AB| = -\frac{13\zeta_1^4 - 62\zeta_1^2 + 60}{6\left(1-\zeta_1^2\right)} \leq 0$$

if and only if $13\zeta_1^4 - 62\zeta_1^2 + 60 \geq 0$. Due to Lemma 1.1 and equations (2.4)-(2.5),

$$|\mathcal{L}| \leq \frac{1}{8}\zeta_1(1-\zeta_1^2)\left(|A| + |B| - |C|\right) = \frac{1}{4} - \frac{\zeta_1^4}{48} \leq \frac{1}{4}.$$

D. Finally the expression below is true for each $\zeta_1 \in X^*$,

$$|AB|-|C|(|B|-4|A|) = \frac{277\zeta_1^4 - 238\zeta_1^2 + 60}{6(1-\zeta_1^2)} > 0.$$

Summarizing **I** and **II** inequality (2.1) follows. In Lemma 1.1, on replacing $\zeta_1 = 0, \zeta_2 = \zeta_3 = 1$ in (1.8) we get $\tilde{p}(z) = (1+z^2)/(1-z^2) \in \mathcal{P}$. The function $\tilde{f} \in \mathcal{A}$ given in (2.2) satisfies $2z\tilde{f}'(z)/(\tilde{f}(z) - \tilde{f}(-z)) = (1+z^2)/(1-z^2)$. Hence the result is sharp. ∎

Theorem 2.2. *Let F_f be given by* (1.3)*. If $f \in \mathcal{F}_2$, then sharp bound on Hankel determinant for F_f is given by*

$$|H_{2,1}(F_f)| \leq \frac{1}{4}. \tag{2.7}$$

Above inequality is sharp due to the function given in Theorem 2.1.

Proof. Since $f \in \mathcal{F}_2$, then there exists an analytic function $p \in \mathcal{P}$ defined in (1.4) such that

$$(1-z^2)f'(z) = p(z). \tag{2.8}$$

Then from (2.8) we obtain first three coefficients of $f(z)$ in terms of $c_i's$ $(i = 1, 2, 3)$ as follows:

$$a_2 = \frac{1}{2}c_1, \quad a_3 = \frac{1}{3}(1+c_2), \quad a_4 = \frac{1}{4}(c_1+c_3).$$

Using Lemma 1.1, we express $\gamma_1\gamma_3 - \gamma_2^2$ in terms of $\zeta_i's$ by replacing $c_i's$ with $\zeta_i's$ $(i = 1, 2, 3)$,

$$\begin{aligned}\mathcal{M} := \gamma_1\gamma_3 - \gamma_2^2 = \frac{1}{144}(&\zeta_1^4(5 - 4\zeta_2 + 2\zeta_2^2) + 2\zeta_1^2(1 + 10\zeta_2 + 7\zeta_2^2) - 4(1+2\zeta_2)^2 \\ &+ 18\zeta_1\zeta_3(1-\zeta_1^2)(1-|\zeta_2|^2))\end{aligned} \tag{2.9}$$

Due to the fact that $|\zeta_2| \leq 1$, inequality (2.9) gives

$$|\mathcal{M}| = \begin{cases} 1/48, & \zeta_1 = 1, \\ (1/36)(1+2\zeta_2)^2 \leq 1/4, & \zeta_1 = 0. \end{cases}$$

Since $|\zeta_3| \leq 1$, then for $\zeta_1 \in (0,1)$, we have

$$|\mathcal{M}| \leq \frac{1}{8}\zeta_1(1-\zeta_1^2)\Psi(A,B,C), \tag{2.10}$$

where

$$\Psi(A,B,C) := |A + B\zeta_2 + C\zeta_2^2| + 1 - |\zeta_2|^2$$

with

$$A = \frac{5\zeta_1^4 + 2\zeta_1^2 - 4}{18\zeta_1\left(1-\zeta_1^2\right)}, \quad B = -\frac{2}{9\zeta_1}(4-\zeta_1^2), \quad C = -\frac{1}{9\zeta_1}(8+\zeta_1^2). \tag{2.11}$$

In light of Theorem 1.2 and the expressions of A, B and C given in (2.11), we divide our result in two cases based on the value of AC.

I. Observe that AC is non-negative over the interval $Y = (0, \zeta^*]$, where $\zeta^* = (1/5)\sqrt{\sqrt{21}-1}$. Since $|B| - 2(1-|C|) = \frac{8}{3\zeta_1} - 2 \geq 0$, for each $\zeta_1 \in Y$, therefore by applying Lemma 1.2 to (2.10) yields the following inequality

$$|\mathcal{M}| \leq \frac{1}{8}\zeta_1(1-\zeta_1^2)(|A| + |B| + |C|) = \frac{1}{48}(12 - 12\zeta_1^2 - \zeta_1^4) := S_1(\zeta_1).$$

It is clearly evident that $S_1(x)$ is a decreasing function of x, so $S_1(x) \to 1/4$ whenever $x \to 0$. Thus for this case $|\mathcal{M}| \leq 1/4$.

II. Let $Y^* = Y^c \cap [0,1)$. Clearly $AC < 0$ for each $\zeta_1 \in Y^*$. For our suitability, we divide the

computations in five subcases.

A. The expression in (2.12) is true for each $\zeta_1 \in Y^*$,

$$S_1(x) := -4AC\left(\frac{1}{C^2}-1\right) - B^2 = \frac{2(\zeta_1^4 - 106\zeta_1^2 - 12)}{27(8+\zeta_1^2)} \leq 0. \tag{2.12}$$

provided $\zeta_1^4 - 106\zeta_1^2 - 12 \leq 0$. Additionally, from Case **I** we have $|B|-2(1-|C|) =: S_2(x) \geq 0$. Thus we conclude that $S_1(Y^*) \cap S_2(Y^*)$ is an empty set.

B. Moreover for $\zeta_1 \in Y^*$, it is easy to verify that

$$\begin{cases} S_3(\zeta_1) := 4(1+|C|)^2 = \dfrac{4}{81\zeta_1^2}(\zeta_1^2+9\zeta_1+8)^2 > 0 \\ \\ S_4(\zeta_1) := -4AC\left(\dfrac{1}{C^2}-1\right) = \dfrac{2(\zeta_1^2-64)(5\zeta_1^4+2\zeta_1^2-4)}{81\zeta_1^2(8+\zeta_1^2)} \leq 0. \end{cases}$$

Therefore $B^2 = 2(\zeta_1^2-4)/9\zeta_1 > \min\{S_3(\zeta_1), S_4(\zeta_1)\} = S_4(\zeta_1)$. This leads us to the next case.

C. Consider the following expression

$$|C|(|B|+4|A|-|AB|) = \frac{S_5(\zeta_1)}{81\zeta_1^3(1-\zeta_1^2)},$$

where $S_5(x) := 2x^8+10x^7+11x^6+84x^5-90x^4+24x^3+52x^2-64x+16$. Observe that for each $x \in Y^*$, the polynomial $S_5(x) \leq 0$ is true if and only if $x^2(2x^6+21x^4+76) \leq 6(8+x^4)$. Infact

$$\min_{x\in Y^*}(x^2(2x^6+21x^4+76)) = 36/625(329\sqrt{21}-419) \geq \max_{x\in Y^*}(6(8+x^4)) = 54.$$

Thus for each $x \in Y^*$, $S_5(x)$ must be positive.

D. Finally the following inequality

$$\begin{aligned} |AB|-|C|(|B|-4|A|) = -\frac{1}{81\zeta_1^3(1-\zeta_1^2)}(&2\zeta_1^8-10\zeta_1^7-5\zeta_1^6-68\zeta_1^5-10\zeta_1^4-104\zeta_1^3 \\ &-12\zeta_1^2+128\zeta_1+16) \leq 0, \end{aligned}$$

is equivalent to

$$\zeta_1^2(10\zeta_1^5+5\zeta_1^4+68\zeta_1^3+10\zeta_1^2+104\zeta_1+12) \leq 2(\zeta_1^8+64\zeta_1+8),$$

provided $\zeta^* < \zeta_1 \leq \sqrt{\sqrt{249}-15}$. Note that the function $R(x)$ defined below is a decreasing function of x $\left(\zeta^* < x \leq \sqrt{\sqrt{249}-15}\right)$. Thus Lemma 1.2 together with (2.10), gives

$$\begin{aligned} |\mathcal{M}| &\leq \frac{1}{8}\zeta_1(1-\zeta_1^2)(-|A|+|B|+|C|) \\ &= \frac{2\zeta_1^5-5\zeta_1^4-34\zeta_1^3-2\zeta_1^2+32\zeta_1+4}{144\zeta_1} =: R(\zeta_1) \\ &\leq \frac{1}{600}(169-29\sqrt{21}) \approx 0.0601755. \end{aligned}$$

E. Further for $\sqrt{\sqrt{249}-15} < \zeta_1 < 1$, inequality (2.10) and Lemma 1.2, yields

$$\begin{aligned}|\mathcal{M}| &\leq \frac{1}{8}\zeta_1(1-\zeta_1^2)\left((|A|+|C|)\sqrt{1-\frac{B^2}{4AC}}\right)\\ &= \frac{1}{16\sqrt{3}}(2-\zeta_1^2)^2\sqrt{\frac{\zeta_1^2\left(\zeta_1^4+20\zeta_1^2-12\right)}{5\zeta_1^6+42\zeta_1^4+12\zeta_1^2-32}}\\ &\leq \frac{1}{8}(3\sqrt{249}-47)\approx 0.0424002.\end{aligned}$$

Based on all the above cases, we arrive at the required bound. Infact on substituting $\zeta_1 = 0, \zeta_2 = \zeta_3 = 1$ in (2.9), equality in (2.7) is achieved. Clearly $\tilde{f} \in \mathcal{A}$ defined in (2.2) belongs to $\mathcal{F}_2$ such that $(1-z^2)\tilde{f}'(z) = (1+z^2)/(1-z^2)$ and it works as the extremal function . ∎

Here below, we provide some Caratheodóry functions that will be relevant for our onward study.

$$p_i(z) = \begin{cases} \dfrac{1+0.0605378z+z^2}{1-z^2}, & i=1\\[2ex] \dfrac{1-z^2}{1-0.747551z+z^2}, & i=2\\[2ex] \dfrac{1+0.261084z+0.261084z^2+z^3}{1-0.501762z+0.501762z^2-z^3}, & i=3. \end{cases} \tag{2.13}$$

Theorem 2.3. *Suppose* $X(x) = -48x^4 - 96x^3 - 392x^2 + 24x + 357$ *and* F_f *be given by* (1.3)*. If* $f \in \mathcal{F}_1$*, then sharp bound on Hankel determinant for* F_f *is given by*

$$|H_{2,1}(F_f)| \leq X(\eta)/2304, \tag{2.14}$$

where $\eta \approx 0.0302$ *is the unique real root of the equation* $X'(x) = 0$*. This result is sharp due to the extremal function*

$$\tilde{f}(z) = \int_0^z p_1(t)/(1-t)dt, \tag{2.15}$$

where $p_1(z)$ *is defined in* (2.13)*.*

Proof. For each $f \in \mathcal{F}_1$, there exists a function $p \in \mathcal{P}$ defined in (1.4) satisfying $(1-z)f'(z) = p(z)$. On comparing coefficients of like power terms, we obtain

$$a_2 = \frac{1}{2}(1+c_1), \quad a_3 = \frac{1}{3}(1+c_1+c_2), \quad a_4 = \frac{1}{4}(1+c_1+c_2+c_3).$$

With the assumption that $c_1 \in [0,2]$, and applying Lemma 1.1, we obtain,

$$\begin{aligned}|\mathcal{N}| := |\gamma_1\gamma_3 - \gamma_2^2| = \frac{1}{2304}|&72(1+2\zeta_1)(1+2\zeta_3(1-|\zeta_2|^2)+2\zeta_1(1+2\zeta_2-\zeta_2^2)+2\zeta_2\\ &+2\zeta_1^3(1-\zeta_2)^2-2\zeta_1^2(\zeta_3(1-|\zeta_2|^2)+\zeta_2-1))+3(1+2\zeta_1)^4\\ &-64(1+2(\zeta_1+\zeta_2)+2\zeta_1^2(1-\zeta_2))^2|.\end{aligned} \tag{2.16}$$

I. In equation (2.16), on substituting $\zeta_1 = 1$, we get $|\mathcal{N}| = 155/2304 \approx 0.067274$.
II. Since $|\zeta_3| \leq 1$, then for $\zeta_1 \in [0,1)$, we obtain

$$|\mathcal{N}| \leq \frac{1}{16}(1+2\zeta_1)(1-\zeta_1^2)\Psi(A,B,C), \tag{2.17}$$

where

$$\Psi(A,B,C) := |A+B\zeta_2+C\zeta_2^2| + 1 - |\zeta_2|^2$$

with

$$A := \frac{11+56\zeta_1-8\zeta_1^2+16\zeta_1^3+80\zeta_1^4}{144\,(1+2\zeta_1)\left(1-\zeta_1^2\right)}, \quad B := \frac{4\zeta_1^2+4\zeta_1-7}{9\,(1+2\zeta_1)}$$

and

$$C := -\frac{2\zeta_1^2+9\zeta_1+16}{9(1+2\zeta_1)}.$$

Based on the range of ζ_1 and the values of A, B and C defined above, we find that $AC < 0$. Further as a consequence of Lemma 1.2, we split our discussions into the following subcases.

A. For $\zeta_1 \in [0,1)$, consider the expression,

$$\mathcal{V}_1 := -4AC\left(\frac{1}{C^2}-1\right) - B^2 = \frac{T(\zeta_1)}{108\,(1+2\zeta_1)\left(2\zeta_1^2+9\zeta_1+16\right)},$$

where $T(x) = 32x^5+336^4-2720^3+856x^2+846x-1687$. It is easy to verify that $\max_{x\in[0,1)} T(x) = -1687$, thus $T(x) \le 0$. Suppose $\zeta^* = (1/2)(2\sqrt{2}-1)$, then for $\zeta_1 \in W_1 \cup W_1^*$, where $W_1 := [0,\zeta^*]$ and $W_1^* := W_1^c \cap [0,1)$, define

$$\mathcal{V}_2 := |B|-2(1-|C|) = \begin{cases} \dfrac{21-22\zeta_1}{18\zeta_1+9}, & \zeta_1 \in W_1, \\ \dfrac{8\zeta_1^2-14\zeta_1+7}{18\zeta_1+9}, & \zeta_1 \in W_1^*. \end{cases}$$

From the expression of $\mathcal{V}_2$, one can observe that, if $\zeta_1 \in W_1$, then $\mathcal{V}_2 \le 0$ leads to $\zeta_1 \ge 21/22$, this is not true. Clearly for $\zeta_1 \in W_1^*$, it is easy to verify that $\mathcal{V}_2 > 0$. Thus $\mathcal{V}_1 \cap \mathcal{V}_2$ is an empty set. Therefore according to Lemma 1.2, this case fails to hold for any value of $\zeta_1 \in [0,1)$.

B. For $\zeta_1 \in [0,1)$, simple computations reveal that

$$\mathcal{V}_3 := 4(1+|C|)^2 = \frac{4(2\zeta_1^2+27\zeta_1+25)^2}{81(1+2\zeta_1)^2} > 0$$

and

$$\mathcal{V}_4 := -4AC\left(\frac{1}{C^2}-1\right) = \frac{(4\zeta_1^2+36\zeta_1-175)(80\zeta_1^4+16\zeta_1^3-8\zeta_1^2+56\zeta_1+11)}{324(2\zeta_1+1)^2(2\zeta_1^2+9\zeta_1+16)} < 0.$$

Above inequalities yield,

$$B^2 = \frac{(7-4\zeta_1-4\zeta_1^2)^2}{81\,(2\zeta_1+1)^{\,2}} > \min\{\mathcal{V}_3,\mathcal{V}_4\} = \mathcal{V}_4.$$

C. Furthermore for $\zeta_1 \in W_1 \cup W_1^*$, we have the following expression

$$\mathcal{V}_5 := |C|\,(|B|+4\,|A|) - |AB| = \begin{cases} \dfrac{Q_1(\zeta_1)}{1296\,(1+2\zeta_1)^2\left(1-\zeta_1^2\right)}, & \zeta_1 \in W_1, \\ \dfrac{Q_2(\zeta_1)}{1296\,(1+2\zeta_1)^2\left(1-\zeta_1^2\right)}, & \zeta_1 \in W_1^* \cup \{\zeta^*\}, \end{cases}$$

where $Q_1(x) := 1088x^6+4096x^5+6352x^4+576x^3-1252x^2+3616x+2419$ and $Q_2(x) := 192x^6+1920x^5+4912x^4+1792x^3+4436x^2+4344x-1011$. A computation reveals that $Q_1'(x)$ is never zero on W_1 and $Q_1'(0) = 3616$, thus $\mathcal{V}_5 \ge Q_1(0)/1296 > 0$. Similarly when $\zeta_1 \in W_1^*$, then $\mathcal{V}_5 \to Q_2(\zeta^*)/(8\sqrt{2}-10) > 0$ as $\zeta_1 \to \zeta^*$. Therefore in view of Lemma 1.2, this case fails to hold

any $\zeta_1 \in [0,1)$.

D. Consider the expression

$$\mathcal{V}_6 := |AB|-|C|(|B|-4|A|) = \frac{Q_2(\zeta_1)}{1296\,(1+2\zeta_1)^{\,2}\left(1-\zeta_1^2\right)}.$$

It can be easily verified that $Q_2'(x)$ is non-vanishing over the range $[0,1)$. Infact $Q_2(0) = -1011$ and $Q_2(1/2) = 2864$, thus by intermediate value property their exists $\gamma \in (0,1/2)$ such that $Q_2(\gamma) = 0$. Infact $Q_2(\gamma) = 0$, provided $\gamma \approx 0.190991$. Thus $Q_2(x) \leq 0$ for each $x \in [0,\gamma]$, this leads to $\mathcal{V}_6 \leq Q_2(\gamma)/(1296(1+2\gamma)^2(1-\gamma^2)) = 0$. Consequently, on applying Lemma 1.2 to inequality (2.17), we have

$$|\mathcal{N}| \leq \frac{1}{16}\,(1+2\zeta_1)\,(1-\zeta_1^2)(|B|+|C|-|A|) \leq \phi(\eta) \quad (\eta \approx 0.0302689),$$

where

$$\phi(x) := \frac{X(x)}{2304} = \frac{-48x^4-96x^3-392x^2+24x+357}{2304} \quad 0 \leq x \leq \gamma.$$

Observe that for $0 \leq x \leq \gamma$,

$$\phi'(x) = \frac{-192x^3-288x^2-784x+24}{2304} = 0,$$

holds true only if $x = \eta < \gamma$, also $\phi''(\eta) < 0$. Consequently, this leads us to the inequality $\phi(x) \leq \phi(\eta) \approx 0.155106$.

E. Furthermore, for $\gamma < \zeta_1 < 1$, from (2.17) and Lemma 1.2, it follows that

$$|\mathcal{N}| \leq \frac{1}{16}\,(1+2\zeta_1)\,(1-\zeta_1^2)(|A|+|C|)\sqrt{1-\frac{B^2}{4AC}} \leq \psi(\gamma) \quad (\gamma \approx 0.190991),$$

where

$$\psi(x) := \frac{1}{768}(48x^4-128x^3-232x^2+200x+267) \times \sqrt{\frac{32x^6+208x^5+544x^4+216x^3+14x^2+257x+124}{480x^6+2256x^5+4224x^4+888x^3+1194x^2+2985x+528}}.$$

As $\psi'(x) = 0$ has no solution in $(\gamma,1)$, also $\psi'(x) \to -(3079/15552) < 0$ when $x \to 1$. Thus $\psi(x)$ is a decreasing function and hence $\psi(x) \leq \psi(\gamma) \approx 0.150413$. In a nutshell, cases **I-II**, leads us to inequality (2.14). In Lemma 1.1, on substituting $\zeta_1 = \eta, \zeta_2 = 1$ and $\zeta_3 = 0$ in (1.8) we obtain $p_1 \in \mathcal{P}$ defined in (2.13). Thus $\tilde{f} \in \mathcal{A}$ defined in (2.15) belongs to $\mathcal{F}_1$ and serves as the extremal function. ∎

Theorem 2.4. *Suppose* $X(x) = -176x^4-224x^3-264x^2+328x+469$. *Let* F_f *be given by* (1.3). *If* $f \in \mathcal{F}_3$, *then sharp bound on Hankel determinant for* F_f *is given by*

$$|H_{2,1}(F_f)| \leq X(\eta)/2304, \tag{2.18}$$

where $\eta \approx 0.3737$ *is the unique real root of the equation* $X'(x) = 0$. *This result is sharp for the function*

$$f(z) = \int_0^z p_2(t)/(1-t+t^2)dt,$$

where $p_2(z)$ *is given by* (2.13).

Proof. Suppose $f \in \mathcal{A}$ such that $(1-z+z^2)f'(z) = p(z)$, where $p \in \mathcal{P}$ is defined in (1.4). Proceeding in the similar approach, we have

$$a_2 = \frac{1}{2}(1+c_1), \quad a_3 = \frac{1}{3}(c_1+c_2), \quad a_4 = \frac{1}{4}(c_2+c_3-1)$$

The expression $\gamma_1\gamma_3-\gamma_2^2$ in terms of $\zeta_i's$ $(i=1,2,3)$, becomes

$$\begin{aligned}\mathcal{G}:=\gamma_1\gamma_3-\gamma_2^2=\frac{1}{32}(1+2\zeta_1)(2\zeta_1^3(1-\zeta_2)^2+2\zeta_1\zeta_2(2-\zeta_2)+2(\zeta_2+\zeta_3)-1\\+2\zeta_1^2(1-\zeta_2-\zeta_3(1-|\zeta_2|^2))-2\zeta_3|\zeta_2|^2)+\frac{1}{768}(1+2\zeta_1)^4\\-\frac{1}{9}(\zeta_1+\zeta_2+(1-\zeta_2)\zeta_1^2)^2.\end{aligned}\tag{2.19}$$

I. In expression (2.19) put $\zeta_1=1$, then $|\mathcal{G}|=133/2304\approx 0.0577257$.
II. If $\zeta_1\in[0,1)$, then from (2.19) together with $|\zeta_3|\leq 1$, gives

$$|\mathcal{G}|\leq\frac{1}{16}(1+2\zeta_1)(1-\zeta_1^2)\Psi(A,B,C),\tag{2.20}$$

where

$$\Psi(A,B,C):=|A+B\zeta_2+C\zeta_2^2|+1-|\zeta_2|^2$$

with

$$A:=\frac{80\zeta_1^4+16\zeta_1^3-40\zeta_1^2-120\zeta_1-69}{144\,(1+2\zeta_1)\,(1-\zeta_1^2)},\quad B:=\frac{4\zeta_1^2+4\zeta_1+9}{9(1+2\zeta_1)}$$

and

$$C:=-\frac{2\zeta_1^2+9\zeta_1+16}{9(1+2\zeta_1)}.$$

The expressions of A,B and C defined above leads to $AC>0$. Now the inequality

$$|B|-2(1-|C|)=\frac{8\zeta_1^2-14\zeta_1+23}{9(1+2\zeta_1)}<0,\tag{2.21}$$

is equivalent to $8\zeta_1^2-14\zeta_1+23<0$. Since $\min_{\zeta_1\in[0,1)}8\zeta_1^2-14\zeta_1+23=135/8>0$, then inequality (2.21) is false for each $\zeta_1\in[0,1)$. Therefore $|B|-2(1-|C|)\geq 0$ for each $0\leq\zeta_1<1$. An application of Lemma 1.2 to inequality (2.20), gives

$$|\mathcal{G}|\leq\frac{1}{16}(1+2\zeta_1)(1-\zeta_1^2)(|A|+|B|+|C|)\leq S(\eta)\quad(\eta\approx 0.373776),\tag{2.22}$$

where

$$S(x):=\frac{X(x)}{2304}=\frac{-176x^4-224x^3-264x^2+328x+469}{2304}.\tag{2.23}$$

Note that $S'(x)$ vanishes at $x=\eta$ and $S''(\eta)<0$, where $\eta\approx 0.373776$ is the only critical point of $S(x)$ in $[0,1)$. Hence (2.22) determines the desired bound. On replacing $\zeta_1=\eta,\zeta_2=-1$ and $\zeta_3=0$ in (1.8) we obtain $p_2\in\mathcal{P}$ defined in (2.13). As a result sharpness in (2.18) is achieved by $\tilde{f}\in\mathcal{A}$ such that $(1-z+z^2)\tilde{f}'(z)=p_2(z)$. ∎

Theorem 2.5. *Suppose* $48(17+x)X(x)=(1+x)(x^4+20x^3-114x^2+4x+125)$. *Let* F_f *be given by* (1.3). *If* $f\in\mathcal{F}_4$, *then sharp bound on Hankel determinant for* F_f *is given by*

$$|H_{2,1}(F_f)|\leq X(\eta),\tag{2.24}$$

where $\eta\approx 0.381$ *is the unique real root of the equation* $X'(x)=0$. *This result is sharp with extremal function*

$$\tilde{f}(z)=\int_0^z p_3(t)/(1-t)^2dt,\tag{2.25}$$

where $p_3(z)$ *is defined in* (2.13).

Proof. Let $f \in \mathcal{A}$ satisfy $(1-z+z^2)f'(z) = p(z)$, where where $p \in \mathcal{P}$ is defined in (1.4). Following the same procedure as before, we have

$$a_2 = \frac{1}{2}(c_1+2), \quad a_3 = \frac{1}{3}(2c_1+c_2+3), \quad a_4 = \frac{1}{4}(3c_1+2c_2+c_3+4).$$

Due to Lemma 1.1, the expression $\gamma_1\gamma_3 - \gamma_2^2$ in terms of $\zeta_i's$ $(i = 1, 2, 3)$, becomes

$$\begin{aligned}\mathcal{H} := \gamma_1\gamma_3 - \gamma_2^2 = \frac{1}{144}(18(1+\zeta_1)^2(\zeta_1^2(1-\zeta_2)^2 - \zeta_1(\zeta_2^2 - 1 + \zeta_3(1-|\zeta_2|^2)) \\ - \zeta_3|\zeta_2|^2 + 2\zeta_2 + \zeta_3 + 2) + 3(1+\zeta_1)^4 - 4(3 + 4\zeta_1 + 2\zeta_2 \\ - 2\zeta_1^2(\zeta_2 - 1))^2). \end{aligned} \tag{2.26}$$

I: Substitute $\zeta_1 = 1$ in (2.26), then $|\mathcal{H}| = 1/12$.
II: Since $|\zeta_3| \in \overline{\mathbb{D}}$, then for $\zeta_1 \in [0, 1)$, we have

$$|\mathcal{H}| \leq \frac{1}{8}(1-\zeta_1)(1+\zeta_1)^2\Psi(A, B, C), \tag{2.27}$$

where

$$\Psi(A, B, C) := |A + B\zeta_2 + C\zeta_2^2| + 1 - |\zeta_2|^2$$

with

$$A := \frac{5\zeta_1^4 + 2\zeta_1^3 - 4\zeta_1^2 + 6\zeta_1 + 3}{18(1-\zeta_1)(1+\zeta_1)^2}, \quad B := -\frac{2(1-\zeta_1)(3+\zeta_1)}{9(1+\zeta_1)}$$

and

$$C := -\frac{1}{9}(8+\zeta_1).$$

The expressions of A, B and C defined above, demonstrates that $AC < 0$. Moreover, based on the conditions in Lemma 1.2, we now look into the following subcases.

A. Note that the inequality

$$|B| - 2(1 - |C|) = \frac{4(1-\zeta_1)}{9(1+\zeta_1)} \leq 0$$

is valid if $\zeta_1 \geq 1$. This is false as $\zeta_1 < 1$. Infact

$$-4AC\left(\frac{1}{C^2} - 1\right) - B^2 = \frac{2(\zeta_1^5 + 21\zeta_1^4 - 10\zeta_1^3 - 2\zeta_1^2 + 93\zeta_1 - 31)}{27(1+\zeta_1)^2(8+\zeta_1)} \leq 0,$$

provided $\zeta_1 \in [0, \zeta^*]$, where $\zeta^* \approx 0.336931$. Consequently, due to Lemma 1.2, this case does not hold for each $\zeta_1 \in [0, 1)$.

B. Note that the following inequalities

$$\mathcal{J}_1 := 4(1 + |C|)^2 = \frac{4}{81}(17+\zeta_1)^2 > 0$$

and

$$\mathcal{J}_2 := -4AC\left(\frac{1}{C^2} - 1\right) = \frac{2(17+\zeta_1)(5\zeta_1^4 + 2\zeta_1^3 - 4\zeta_1^2 + 6\zeta_1 + 3)}{81(1+\zeta_1)^2(8+\zeta_1)} > 0,$$

hold simultaneously if $\zeta_1 > \zeta^*$. Thus

$$B^2 = \frac{4(1-\zeta_1)^2(3+\zeta_1)^2}{81(1+\zeta_1)^2} < \min\{\mathcal{J}_1, \mathcal{J}_2\} = \mathcal{J}_2.$$

Therefore in view of Lemma 1.2, inequality (2.27), gives

$$|\mathcal{H}| \leq \frac{1}{8}(1-\zeta_1)(1+\zeta_1)^2\left(1 + |A| + \frac{B^2}{4(1+|C|)}\right) \leq X(\eta) \quad (\eta \approx 0.381423),$$

where

$$X(x)=\frac{(1+x)(x^4+20x^3-114x^2+4x+125)}{48(17+x)}.$$

Since $x=\eta>\zeta^*$ is a unique real root of $X'(x)=0$ in $(\zeta^*,1)$ and $X''(\eta)<0$, where $\eta\approx 0.381423$. Then X attains its maximum at $x=\eta$.

C. Further for $0\le\zeta_1\le\zeta^*$, simple computations reveal that

$$\begin{aligned}|C|(|B|+4|A|)-|AB| &= \frac{17\zeta_1^6+128\zeta_1^5+137\zeta_1^4-80\zeta_1^3-17\zeta_1^2+160\zeta_1+87}{81\,(1-\zeta_1)\,(1+\zeta_1)^{\,3}}\\ &\ge m>0\quad (m\approx 1.05323).\end{aligned}$$

Similarly $|AB|-|C|\,(|B|-4\,|A|)>0$. Therefore, Lemma 1.2, together with (2.27), establishes the following

$$\begin{aligned}|\mathcal{H}| &\le \frac{1}{48}(3\zeta_1^4-16\zeta_1^3-18\zeta_1^2+24\zeta_1+19)\\ &\quad\times\sqrt{\frac{\zeta_1^5+12\zeta_1^4+8\zeta_1^3-2\zeta_1^2+3\zeta_1+14}{15\zeta_1^5+126\zeta_1^4+36\zeta_1^3-78\zeta_1^2+153\zeta_1+72}}\le n\quad (n\approx 0.183792).\end{aligned}$$

By summarizing all the above cases, the required bound can be established. Here equality in (2.24) holds if we replace $\zeta_1=\eta,\zeta_2\approx -0.0871127$ and $\zeta_3=1$ in (2.26). On substituting these values of $\zeta_i\,'s$ $(i=1,2,3)$ in (1.8) we obtain $p_3\in\mathcal{P}$ defined in (2.13). Infact $\tilde{f}\in\mathcal{A}$ given in (2.25) satisfies $(1-z)^2\tilde{f}'(z)=p_3(z)$. Hence $\tilde{f}\in\mathcal{F}_4$ serves as the extremal function. ∎

References

[1] Ali, M.F., Vasudevarao, A.: On logarithmic coefficients of some close-to-convex functions. Proc. Am. Math. Soc. **146**, 1131–1142 (2018)

[2] Cantor, D.G.: Power series with integral coefficients. Bull. Amer. Math. Soc. **69**, 362–366 (1963)

[3] Carathéodory, C.: Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen. Math. Ann. **64**, 95–115 (1907)

[4] Cho, N.E., Kowalczyk, B., Kwon, O.S., Lecko, A., Sim, Y.J.: On the third logarithmic coefficient in some subclasses of close-to-convex functions. Rev. R. Acad. Cienc. Exactas Fís. Nat. **114**(52), 1—14 (2020)

[5] Cho, N.E., Kowalczyk B., Lecko, A.: Sharp bounds of some coefficient functionals over the class of functions convex in the direction of the imaginary axis. Bull. Aust. Math. Soc. **100**(1), 86–96 (2019)

[6] Choi, J.H., Kim Y.C., Sugawa, T.: A general approach to the Fekete-Szegö problem. J. Math. Soc. Jpn. **59**(3), 707–727 (2007)

[7] Dienes, P.: The Taylor series: an introduction to the theory of functions of a complex variable. Dover Publications, Inc., New York (1957)

[8] Ehrenborg, R.: The Hankel determinant of exponential polynomials. Amer. Math. Monthly. **107**(6), 557–560 (2000)

[9] Goel, P., Kumar, S.: On sharp bounds of certain close-to-convex functions, arXiv:1907.13385v2

[10] Goodman, A.W.: Univalent functions. Vol. I, Mariner Publishing Co., Inc., Tampa, FL (1983)

[11] Goodman, A.W.: Univalent functions. Vol. II, Mariner Publishing Co., Inc., Tampa, FL (1983)

[12] Hayman, W.K.: On the second Hankel determinant of mean univalent functions. Proc. London Math. Soc. **18**, 77–94 (1968)

[13] Janteng, A., Halim, S.A., Darus, M.: Coefficient inequality for a function whose derivative has a positive real part. JIPAM. J. Inequal. Pure Appl. Math. **7**(2), 1–5 (2006)

[14] Kaplan, W.: Close-to-convex schlicht functions. Michigan Math. J. **1**, 169–185 (1952)

[15] Kowalczyk B., Lecko, A.: The Fekete-Szegö inequality for close-to-convex functions with respect to a certain starlike function dependent on a real parameter. J. Inequal. Appl. **2014**, 1–16 (2014)

[16] Kowalczyk, B., Lecko, A.: Second Hankel determinant of logarithmic coefficients of convex and starlike functions. Bull. Aust. Math. Soc. **105**(3), 458-467 (2022)

[17] Kumar, U.P., Vasudevarao, A.: Logarithmic coefficients for certain subclasses of close-to-convex functions. Monatsh. Math. **187**(3), 543–563 (2018)

[18] Kumar, V.: Moduli difference of successive inverse coefficients for certain classes of close-to-convex functions. Ricerche di Matematica (2021). https://doi.org/10.1007/s11587-021-00682-1

[19] Kumar, S., Kumar, V.: Sharp estimates on the third order Hermitian-Toeplitz Determinant for Sakaguchi Classes. Commun. Korean Math. Soc. (2021). https://doi.org/10.4134/CKMS.c210332
[20] Hayman, J.W.: The Hankel transform and some of its properties. J. Integer Seq. **4**, 1-11 (2001)
[21] Lee, S.K., Ravichandran V., Supramaniam, S.: Bounds for the second Hankel determinant of certain univalent functions. J. Inequal. Appl. **2013**, 281-297 (2013)
[22] Libera, R.J., Złotkiewicz, E.J.: Early coefficients of the inverse of a regular convex function. Proc. Amer. Math. Soc. **85**(2), 225–230 (1982)
[23] Libera, R.J., Złotkiewicz, E.J.: Coefficient bounds for the inverse of a function with derivative in $\mathcal{P}$. Proc. Amer. Math. Soc. **87**(2), 251–257 (1983)
[24] Milin, I.M.: Univalent Functions and Orthonormal Systems, Izdat. "Nauka", Moscow (1971) (in Russian); English transl., American Mathematical Society, Providence (1977)
[25] Noonan, J.W., Thomas, D.K.: On the second Hankel determinant of areally mean p-valent functions. Trans. Amer. Math. Soc. **223**, 337–346 (1976)
[26] Noor, K.I.: Higher order close-to-convex functions. Math. Jpn. **37**(1), 1–8 (1992)
[27] Ozaki, S.: On the theory of multivalent functions. II. Sci. Rep. Tokyo Bunrika Daigaku. Sect. A. **4**, 45–87 (1941)
[28] Pommerenke, C.: On the coefficients and Hankel determinants of univalent functions. J. Lond. Math. Soc. **41**, 111–122 (1966)
[29] Pommerenke, C.: Univalent Functions. Vandenhoeck & Ruprecht, Göttingen (1975)
[30] Sakaguchi, K.: On a certain univalent mapping. J. Math. Soc. Jpn. **11**, 72–75 (1959)
[31] Thomas, D.K.: On the logarithmic coefficients of close to convex functions. Proc. Amer. Math. Soc. **144**(4), 1681–1687 (2016)
[32] Thomas, D.K., Tuneski, N., Vasudevarao, A.: Univalent Functions: A Primer. De Gruyter Studies in Mathematics, **69**, De Gruyter, Berlin, Boston (2018)

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: mridulamundalia@yahoo.co.in

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

Comparative Analysis of Different hand gestures Detection and Recognition Techniques

Rajat Kumar
Department of Information Technology
Delhi Technological University
Delhi, India
rajatkumar_2k18it093@dtu.ac.in

Puneet Sharma
Department of Information Technology
Delhi Technological University
Delhi, India
puneetsharma_2k18it092@dtu.ac.in

Dr. Bindu Verma
Department of Information Technology
Delhi Technological University
Delhi, India
binduverma@dtu.ac.in

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835824

***Abstract*—Hand Gesture is a wide technology aspect of computer Science. It provides a good technology tool for human computer interaction. With some mathematical algorithms, Humans can interact with computers without using any mechanical devices. Also for deaf and mute people, communication hindrance can be remediate. Hand gesture is the most expressive and frequent among any type of expression language. It can put up sign language recognition to a new level. This paper is oriented towards a comparative analysis of different types of hand gesture Detection and recognition methods. In addition, the result and problems with different approaches are tabulated. These have different gesture types and so different number of gesture classes. Most of these methods produce accuracy of 90-95%. The best accuracy achieved for static hand gestures is 95.7% which is based on the ML algorithm.**

Keywords*:- *HSV (Hue, Saturation and value), Hand Detection, Hand Recognition, CNN(Convolutional Neural Network), Convex Hall, K curvature, distance transform

Introduction

Gesture is the way of expressing your emotions with actions. Body and hand gestures indicate a sign .The movement of hand and body convey some message. It provides a method for interacting with machines. In today's era technology has taken great control over us. With the help of technology human's life has become simpler and faster. Computers and Computing Devices are becoming an essential part of human's life. Research in Human Computer Interaction has given a tremendous boost in computer vision. Computer Human Interaction plays a vital role in computer vision. It started off with keyboard, mouse and then voice and speech recognition. Gestures are the most popular and trending way of communication because these all are easy to understand and communicative motions that help to communicate to the environment. Gestures are decreasing the requirement of mechanical devices. Gestures can be of any type like hand, finger ,facial gestures, or whole body gestures like eating, sleeping and walking .

Gestures can be further classified in two parts: static and Dynamic gestures. Static Gesture image is a single image that can be easily processed but dynamic gestures contains movement of the hand that performs the gesture. Hand gestures are a popular field such as medical where the application named gesture makes the doctor study the X ray, CT scan and MRI's. In automotive, the BMW 7 Series has an HCI System which allows the user to control music player , incoming and outgoing calls and much more. There are many more applications like robot controllers, Sign Language recognition systems, Augmented Reality(AR), TV applications, mobile applications, intelligent vehicle [15] and much more. In hand recognition technology, the camera captures the user body movement image and transfers it to the computer. These recognised gestures are used to control the operating system of gadgets. virtual mouse and keyboard are most popular applications of human computer interaction. In [10-12] the author described the procedure of hand gesture recognition. First, the author identifies the hand part from the source image with some input devices. Then features from it are captured to analyse the hand gesture. Hand gesture is recognised by comparing it with featured data. The input device such as stereo camera is used for extracting the hand images and other information. Stereo camera and tof provides the depth information which help to separate the hand region with background of image. Similarly, author in paper [13] used template matching to recognise the NTU static hand gestures.

In this paper we compare some hand gesture techniques in terms of their procedure ,lightning condition , dataset, models used in them and provide a table of their accuracy and with some limitations and propose a better and latest method in order to do hand gesture detection and recognition.

Literature survey

A. Real Time Hand Gesture Recognition Using Finger Segmentation

This method is one of the useful methods for hand gesture recognition and detection. This method is divided into four major parts: **Hand Detection, Finger and Palm segmentation, Finger Recognition and Hand gesture recognition.**

In hand detection simple images(captured by normal camera) with the same background are used. all images have identical backgrounds so it is easy to detect the hand region using the background subtraction method. The skin colour is used to differentiate the hand part from other moving objects. HSV (Hue, Saturation and value) Model is used to detect the skin colour and has the value of 315,94 and 37.

In Finger and palm Detection , a binary image is obtained from the hand detection by background subtraction with white pixels that belong to hand region and black for background. After that, it processes with following procedure:

Palm point(Centre point of palm) is found by using the distance transform methods. In the distance transform

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:26 UTC from IEEE Xplore. Restrictions apply.

image every pixel in the hand region calculates distance with nearest or the boundary pixels having white pixel. The pixel with the greatest distance value among all the white region pixels is considered the palm point as shown in fig 1.

Inner Circle of maximal Radius, when the palm point is identified, using that palm point we draw a circle inscribed in palm and it is known as inner circle and the circle's radius steadily expands until the circle does not touch the palm edges. The final circle is known as the inner circle of maximal radius.

Wrist point and palm mask, the inner circle's radius is obtained then a new circle is drawn with a radius of 1.2 times of inner circle radius and having the blue color and the two wrist points are two terminal points in the hand as shown in fig 1 (figure reprinted from [6])

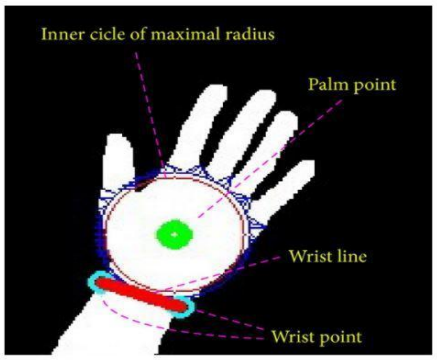


Fig 1 Hand Segmentation

Hand Rotation: After obtaining the palm point and wrist point, then it can produce an arrow which is in the middle of wrist point and the arrow is set to north. When the hand is moved synchronously makes the gestures invariant. The part below the wrist line is cut down to make the hand image more accurate.

Finger Recognition, with help of palm mask, finger, and palm could be properly classified while the other parts of hand would be fingers. The result of segmentation is given in fig 2 (a-d) (figures reprinted from [6]).

Fig. 2(a) The Minimal Bounding Box

Fig .2(b)The Segmented Fingers

Fig. 2(c) The palm Line

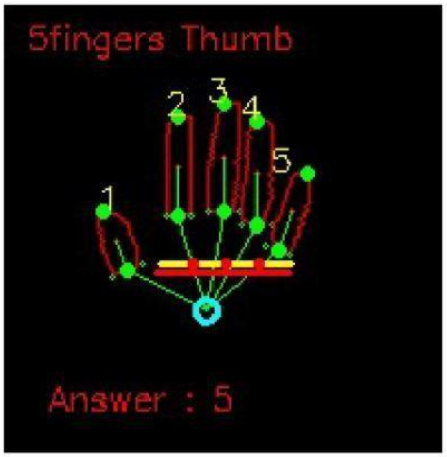


Fig. 2(d)The Recognition of Fingers

In Finger recognition, the segmentation of fingers is based on the labelling algorithm, where the pixels are low in that area considered as a noisy area and that would be discarded and the region with enough size is fingers and remains. It includes two parts: **Thumb detection and Recognition, Detection and recognition of the other fingers.**

In recognition of hand Gesture, after detecting the hand and fingers we use the simple rule classifiers in order to find the Hand gestures. In simple use classifiers, the hand gesture is found using the content of fingers and hand. It is very effective and efficient[6].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:26 UTC from IEEE Xplore. Restrictions apply.

B. *Real Time Hand Gesture Recognition in Depth image using CNN*

Author used a method which is totally based on the CNN architecture model. The benefit of using CNN is that it detects the features automatically, decreasing human supervision. This method is mainly designed to reform communication barrier for deaf and dumb people. The proposed model is able to recognise 11 different static hand gestures by depth images with a very good accuracy. It uses the Senz3d dataset depth images. . The dataset consists of 1320 sample images.

Here the steps involved are: Acquiring depth image, Hand segmentation, Pre-processing, Binary Image conversion, Train CNN model, Fit the training set to CNN and at last Hand Gesture Recognition. Firstly, the hand region is extracted from depth images (taken from dataset). There is a predefined threshold value. The YCbCr value of every pixel is compared with the standard value(i.e., 0 <=Y<=255 , 135 <=C<=180 and 85<=Cb<=135) .
For pre-processing, we resize or crop these segmented images to 256 x 256.
As binary Image is easier to process than multiple color channels, binary image conversion is done using threshold transforms. Pixel value more than 60 threshold value gets converted to 1 and less than it gets converted to 0.

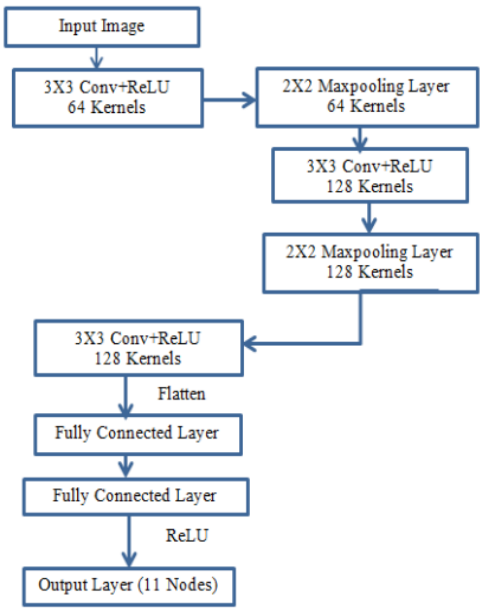


Fig 3: Architecture of proposed CNN model

In this proposed CNN architecture as shown in above fig.3 (figure reprinted from paper [7]) , we have three convolution layers, two max pooling layers and the two fully connected layers as shown in figure. There are 11 nodes in the output layer to recognize 11 gestures. We also have a ReLU activation function that determines the probability for classifying labels. If it is positive, this function directly outputs that input part else output would be zero. For all 11 gestures, a good level of accuracy is achieved[7].

C. *Static Hand Gesture Recognition Using Artificial Neural Network*

An ANN is a system made up of a large number of interconnected nodes. A neural network model links input data sets to expected output datasets and then passes through one or more hidden Layers. In this system, ANN is used with feed forward and backward propagation techniques to train a network with specific characteristics to correctly recognise specific movements. An example of ANN is shown in figure 4 (image reprinted from paper-24. 10.1007/s00500-020-04860-5. *2020)*

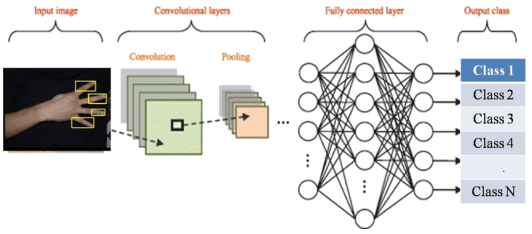


Fig 4. Artificial Neural Network

The recognition of static hand gestures is separated into four parts: Image capturing, Image Processing, Features extraction and Classifications.
Image Capturing: An external or built-in camera can capture a real-world image which is having a view of the person's hand performing the gestures. the frame of each image will be incorporated into the overall flow of the process. RGB, Gray, and HSV colour spaces are utilized to capture images. The image is configured to a 20× 20 resolution with a set of 400pixel for every image when it is captured using the RGB colour space model.
Processing of Real Time Images: Image processing is most important in order to obtain a desired result from the capture of a webcam. Firstly Using a threshold method, we partition the backdrop with the hand. According to the HSV Colour value of the identified object, a range must be specifically provided. Using Gaussians blur, the image would be blurred.
Features extraction: After that, in order for the system to assess the image, the separation of objects from the surrounding contours must be found. There may be more than one contour present, which is why the contour which is of large size is important. For achieving the largest contour, we must iterate over the whole contour and retrieve it using a straightforward method of determining the largest area.
Classification: Neural Networks don't take real images in the input layer. Firstly these images are pre-processed and converted to binary images. a pixel with white colour will be assigned with 0 value and black would have value 1. The layer of input directly depends upon the number of every image's pixels. Therefore , the input layer is set to have 400 pixels for each image. The count of neurons in the output layer is 39 which includes the total number of 10 numbers.26 alphabet , 3 punctuation marks . The mean between the input and output layers is used to determine the number of hidden layers[8].
Author verma et.al in paper [14] used convolutional neural network to extract the features of dynamic hand gesture and classification is performed using Grassmann discriminant analysis to classify the gestures.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:26 UTC from IEEE Xplore. Restrictions apply.

D. *Hand Gesture Recognition using Neural Networks and K-Curvature with convex hall*

Here CNN is used for hand detection and for extracting the features, the algorithm proposed is simply K curvature's combination and the convex hull algorithms. It mainly detects the fingertips. The features can be fingertip finder, pixel segmentation, eccentricity, elongated ness and rotation. Here steps involved are: Image Acquisition, Pre-processing, Feature abstraction, Feature vector, ANN, Classifying gestures.

First to detect hands in image, it is resized to low resolution, and then transformed from RGB to binary using the Global histogram threshold. Now it is passed through the trained neural networks. In the neural network's training phase, the input vector is feature vectors, and the target vector is the number of signs that are used for recognition. After the detection of hand gestures, now we need to extract the features. Convex hull creates a polygon around the segmented hand region, enclosing all of the region's points. We also have our detected hand gesture through CNN. To find the common borders, we perform AND operation between these two to obtain fingertips. Using finger detection, we can extract other characteristics such as the distance between the fingertips and the centroid, rotation which might be useful in detecting various hand gestures[2].

Table 1. Advantages and limitation of above hand gesture detection techniques

Paper	Algorithm used	Advantage	Limitations
[6]	Using finger detection method	Fast hand gesture recognition (average time of 0.02) and good accuracy	with moving objects and objects having colour similar to skin colour degrades the performance
[8]	Using artificial neural Networks	Many hand gestures considered for all alphanumeric characters and few others	Inadequate lighting has an impact on accuracy because the colour value of hand gloves is insufficient.
[9]	Using Neural Networks and K convex Hall	combines multiple algorithms so performs better in terms of accuracy and recognized classes	Features-Eccentricity and elongated ness provides low accuracy
[7]	Using CNN trained with depth image	Good level of accuracy for all gestures	Less gestures considered, (0-10 finger counts)

Proposed work

Hand Gesture Recognition using media pipe library

In this Method we utilize the Ml pipeline mediapipe which contains multiple ML Models. Full images are processed with palm detection model and extracts the image oriented box. Hand landmark model is used for cropped image regions extracted with the help of palm detectors and return high precise 3d Hand key points. Processing the cropped image to hand landmark model degrades the requirement of data augmentation. Instead, the network may spend the entirety of its resource to increase coordinate prediction accuracy. Pipeline crops can be developed based on the hand landmarks identified in the previous frame, and palm detection is only there to re-localize the hand when the landmark model can no longer detect its presence.

A.Palm Detection Model

This model helps to detect the location of the hand. This model is designed in such a way as a single shot Detector model detecting the hand. As we know detecting the hand is a difficult task .This model includes two other models: The lite and full models can identify occluded and self-occluded hands and work across a wide range of hand sizes with a significant scale span (20x) compared to the image frame. Faces include unique complexity, such as the eye and mouth; nevertheless, the lack of such complexities in the case of hands makes it more difficult to distinguish them only on the basis of their visual features.

This model includes some strategies, firstly in this method, Because estimating the bounding box of palms and fists is easier than recognising hands with fingers, a palm detector is trained instead of a hand detector. Because palms are smaller, we can easily apply non-maximum suppression methods, which are ideal for two-hand self-occlusion scenarios such as handshakes. Palms are designed with square bounding boxes and reduce the numbers of anchors by the factors of 3-5. And there is a focus loss during training. We minimize this loss to sustain a large number of anchors.

B. Hand Landmark Model

After the palm detection over the full images of the hand , The Hand Landmark Model uses regression to obtain the precise key point locations of 21 3d hand knuckle Coordinates within the observed hand region.. Even with partially visible hands and self-occlusions, the model develops a consistent internal hand posture representation. We labelled 30k real-world photos with 21 3d Coordinates to acquire ground truth data. As demonstrated in Figure 5 (mediapipe documentation, github) , this model creates a high quality synthetic hand model over multiple image backgrounds and maps it to its 3d coordinates in order to cover better hand positions and provide better supervision on the hand geometry.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:26 UTC from IEEE Xplore. Restrictions apply.

Fig.5 Hand Landmark

We can use the mediapipe pipeline to detect various hand gestures. These hand gestures can be recognised for sign language detection and we can also command our computer to perform certain functions using these hand gestures. For example, in case of mediaplayer, we can use the hand gestures to turn volume up/down, rewind etc. as the mediapipe is much accurate and fast for hand gesture detection.

Result and Comparative Study

Table 2. Comparing different hand gesture detection methods

Paper	Algorithm Used	No. of Classes	Result
[6]	Using finger detection method	-	classification accuracy of 1300 images is 92.69 %
[8]	Using artificial neural Networks	39	In real time, accuracy of 90% for a distance of 50cm, 64% for 100cm
[9]	Using Neural Networks and K convex Hall	37	In real time, accuracy was 94.32%.
[7]	Using CNN trained with depth image	11	In real time, accuracy was 94.61%
	Proposed method (using Mediapipe library)	-	average accuracy of 95.7%

This method is tested over real-world images and we got a good accuracy around 95.7% with different lighting conditions and different backgrounds. Also it is much faster than other methods (20 fps average) since it uses the GPU acceleration and multithreading.

CONCLUSION

Hand gesture recognition can be a better method to interact with real world applications as controlling through hands seems more natural and simple. In order to detect and recognize the gestures, there are multiple methods available in this domain. In this paper, we compared different real time hand gesture recognition methods. They all have quite good accuracy but also have few limitations. In case of ML library, Media pipe for recognition of hand gestures which gives an exceptional accuracy.

REFERENCES

[1] Dr. Jayshree Pansare, Prof. Gajanan Aochar, Tejashree Salvi, Jessica Braganza,Akanksha Tiwari, Damini Kesharkar,"Effective Computer Vision Techniques for Real-Time Hand Gesture Recognition and Detection"

[2] Okan Kop¨ ukl ¨ u¨ , Ahmet Gunduz , Neslihan Kose , Gerhard Rigoll "Real-time Hand Gesture Detection and Classification Using Convolutional Neural Networks"

[3] Jing-Hao-sun,Ting Ting Ji,Shu-Bin zhan and jian-kui-Yang "Real time hand gesture Recognition Based on Deep learning"

[4] Dinh-Son Tran, Ngoc-Huynh Ho, Hyung-Jeong Yang ,Eu-Tteum Baek , Soo-Hyung Kim and Gueesang Lee (2020) "Real-Time Hand Gesture Spotting and Recognition Using RGB-D Camera and 3D Convolutional Neural Network"

[5] Sarang Suresh Kakkot, Saylee Gharge (2018) ""Real Time Hand Gesture Recognition & Its Applications In Assistive Technologies For Disabled"

[6] Zhi-hua Chen,1 Jung-Tae Kim,1 Jianning Liang,1 Jing Zhang,1,2 and Yu-Bo Yuan1 "Real-Time Hand Gesture Recognition Using Finger Segmentation "

[7] Boshir Ahmed, Sanchita Rani Das, Dardina Tasmere (2021) "Real Time Hand Gesture Recognition in Depth Image using CNN"

[8] Rosalina, Lita Yusnita, Nur Hadisukmana, R.B Wahyu, Rusdianto Roestam,Yuyu Wahyu ""Implementation of Real-Time Static Hand Gesture Recognition using Artificial NeuralNetwork "

[9] Md. Mohiminul islam, Sarah Siddiqqua and javada afnan3(2017) "Real time hand gesture Detection and recognition using different algorithms based on american sign Lanugage"

[10] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using kinect sensor," IEEE Transactions on Multimedia, vol. 15, no. 5, pp. 1110–1120, 2013.

[11] Z. Mo and U. Neumann, "Real-time hand pose recognition using low-resolution depth images," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06), pp. 1499–1505, June 2006.

[12] S. Miyamoto, T. Matsuo, N. Shimada, and Y. Shirai, "Real-time and precise 3-D hand posture estimation based on classification tree trained with variations of appearances," in Proceedings of the 21st International Conference on Pattern Recognition (ICPR '12), pp. 453–456, November 2012.

[13]. B. Verma, and A. Choudhary. "Unsupervised learning based static hand gesture recognition from RGB-D sensor." In *International Conference on Soft Computing and Pattern Recognition*, pp. 304-314. Springer, Cham, 2016.

[14]. Verma, B., & Choudhary, A. (2018, December). Dynamic Hand Gesture Recognition using Convolutional Neural Network with RGB-D Fusion. In *Proceedings of the 11th Indian Conference on Computer Vision, Graphics and Image Processing* (pp. 1-8).

[15]. Verma, B., & Choudhary, A. (2018). Framework for dynamic hand gesture recognition using Grassmann manifold for intelligent vehicles. *IET Intelligent Transport Systems*, *12*(7), 721-729.

[16] Verma, B., & Choudhary, A. (2020). Grassmann manifold based dynamic hand gesture recognition using depth data. *Multimedia Tools and Applications*, *79*(3), 2213-2237.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:26 UTC from IEEE Xplore. Restrictions apply.

Deep learning based Novel Vehicle Detection and Counting System

Varsha Sisaudia
Department of Information Technology
Delhi Technological University
Delhi, India

Sachin Kumar Giri
Department Of Information Technology
Delhi Technological University
Delhi, India
sachinkumar_bt2k18@dtu.ac.in

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835960

***Abstract*—When researchers talk about smart cities, management of traffic on the highway and road tracking of vehicles is very difficult but in the last few decades, technology has played a great role in solving this problem of traffic management on roads and highways. In this paper, we are going to discuss a similar kind of technology. This paper represents an identification technique based on VISION, which means tracking through cameras. It is mainly targeted at detecting vehicles, counting total number of vehicles, and detecting speed of every vehicle inside the frame.**

Keywords—Deep Learning Algorithms, OPEN_CV library, Visual studio code, python, Captured Traffic Video, yolo4 algorithm, GMM technique.

I. INTRODUCTION

In the last few decades, we can see that there has been a rapid growth in number of vehicles running on roads all through the world. This has led to increase in number of accidents globally. Hence, it is critical to manage traffic smartly and limit the speed of vehicles, especially in accident prone areas or zones. Some tools are used for finding the speed of vehicles but these tools can be inaccurate in some cases. Also, it's difficult for these tools to detect vehicles changing in speeds too often or too fast. Therefore, there is a need for better technology that can be used for detecting the speed of moving vehicles. For this problem, rather than using expensive tools like radar sensors [1], we can use streaming or recorded video of vehicles to solve this problem. The video of moving vehicles will be used as an input and then the videowill be passed through the various kinds of filters inside the OpenCV tool [2] to detect the vehicle, count the vehicle, and find the speed of every vehicle. there are two segments- the first segment is to detect the vehicle and count it whereas the second segment is used to find speed of vehicle when it is in moving position. Fig. 1 shows the flowchart for vehicle categorization and counting.

II. RELATED WORK

For many years, computer vision researchers have been working on moving automobiles in video streams For tracking cars in crowded traffic situations, real-time systems are defined in this paper which is based on feature-based techniques. When we talk about handling tracking, we are talking of categorizing all the sub-types of vehicles. The backgrounds are designed as a slow time-varying sequence that can adjust to variations in lighting and weather. Another well-known High-angle video is used in traffic counting systems.

In this paper videos of traffic are collected by us from online cameras and these videos are used for training of image data and classification and detection of vehicle.

In the case of the night when the area surrounding the vehicle is dark, because of unavailability of street lights, drivers switch on their headlights. The ground is illuminated by their headlights, and this area is considered a moving object. High-intensity area detection and categorization for autos and motorcycles are included in this headlight detection approach.

This research on vehicle detection, tracking, counting vehicle categorization, and detecting the speed of vehicles in this paper vehicles categorization has received very little attention. This is due to the fact that vehicle classification is a difficult task. The identification and monitoring of vehicles are only the beginning of the classification process.

In this paper, for counting vehicles, we use OpenCV tool which is very helpful for maintaining count a vehicle that is moving in the video frame. To store the types of vehicles and counts of vehicles, we make a dataset. This dataset is very helpful for traffic management and helps keeping a track of what kind of vehicles are more prone to accidents in which zones/areas or weather condition.

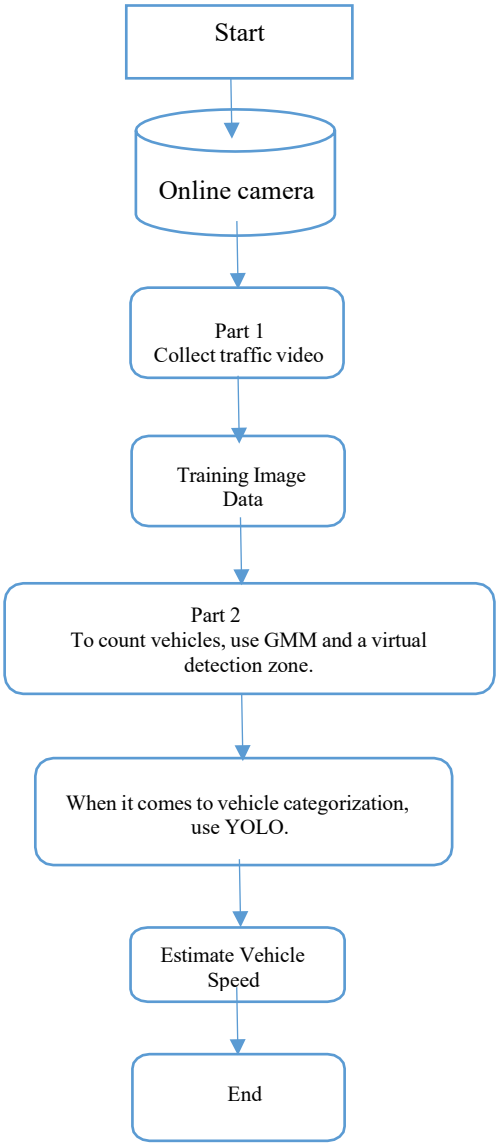


Fig 1. Flowchart for the vehicle categorization and counting procedure

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:02 UTC from IEEE Xplore. Restrictions apply.

III. METHDOLOGIES

1. Input Video:

Our important need is to get the live streaming of moving vehicles using a camera. For this purpose, we have used OpenCV. The video captured through the camera is converted into grayscale for further processing. A video capture object is used for creating a live stream video.

2. Background Subtraction:

For background subtraction [3], we have to collect the video from cameras and this video will be utilized to train picture data as well as classify and verify vehicles

3. Preprocessing the video:

The idea behind pre-processing the video is for improving the information of pictures and images in a video. The number of subcomponents that use an input picture to apply different repairs or enhancement characteristics [4]. If you enable one or more of the preprocessing options, The picture is adjusted using the subcomponents.

4. Dataset preparation:

The dataset used in this paper was collected through the recorded videos of cameras. In India, traffic footage is captured using mounted cameras along various highways. The traffic footage was used to extract visual data, which was then labelled using open-source software called "labelling"[5]. The Ministry of Roads categorizes automobiles based on their common categories. Transport of highways has commonly divided vehicles into six segments. This paper also divides the vehicles into six segments based on different sizes, such as trucks, bikes, cars. in the training process, we have classified and listed the six different segments in a list along with the image in Table 1.

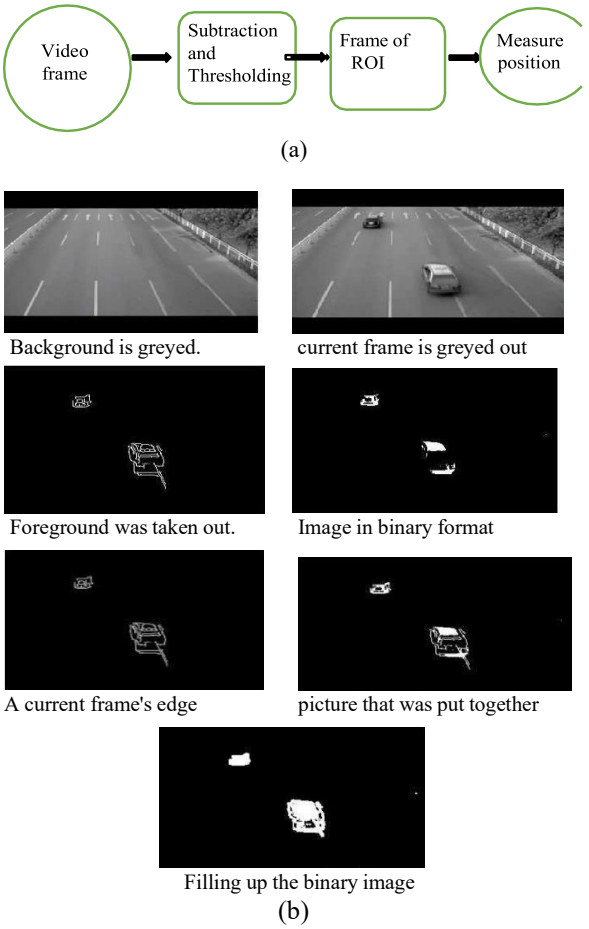


Fig. 2 (a) Background Subtraction (b) Preprocessing the video frame

5. Vehicle Detection and classification:

In this paper, we have used an algorithm that is mainly used for classifying objects called the You Only Look Once (YOLO) algorithm. The YOLO algorithm classifies the vehicles into six segments. Videos are gathered as a validation approach for checking the vehicle categorization. A classifier which is mainly working on YOLO techniques is utilized. Fig. 3 shows a vehicle belonging to one of the six classes is spotted during the training process. The bounding boxes are extracted, their classes are manually labelled, and the information is then submitted to the YOLO model for classification. This study made use of the YOLO4 model art architecture seen in the picture. Darknet 53[6] received images at a resolution of 416 * 416 pixels. This network of feature extraction comprises of 53 convolution layers. It is called as Darknet 53 network.

Darknet employs 53 network kernels, with a batch normalizing layer applied after each convolution layer for normalization. Resnet is a feature extraction method employed by the YOLO algorithm, and then it is used for feature extraction.

For each class, use the Sigmoid function [7]. When a class's score exceeds the threshold, the object is considered to be successful. is consider as this belongs to category one of them by using this function simultaneously another object is also able to identify the category in between multiple classes without any conflict.

The "You Only Look Once" (YOLO) algorithm is popular because it achieves high accuracy while running in real-time. In the sense that it only requires one forward propagation pass through the network to create predictions, this method "only glances once" at the image. It then outputs recognised items together with bounding boxes after non-max suppression

6. Speed Detection

In this work, the real time vehicle speed is also computed. The y axis represents the length of automobile, while the x axis represents the width. Yellow line (referred to as L) in the red circle, according to the scaling in the video, has a length of 4mper Indian traffic standards. A GMM[8] is used to draw a virtual detection zone (blue box) on the route that will be evaluated (referred to as Q). The green box represents the car's frame (referred to as C), and Ct is the car's halfway.

The technique determines the parallel L line segment of Ct (referred to as L) by computing the L distance y between A and B. It then stores y as a consequence of its scale link with.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:02 UTC from IEEE Xplore. Restrictions apply.

TABLE 1.

Class	Vehicle	Length(m)	Image
1.	Sedan	3.5-5.5	
2.	Truck	5.5-11	
3.	Scooter	1-2.5	
4.	Bus	7-12	
5.	Flikcar	15-18	
6.	Hlinkcar	15-18	

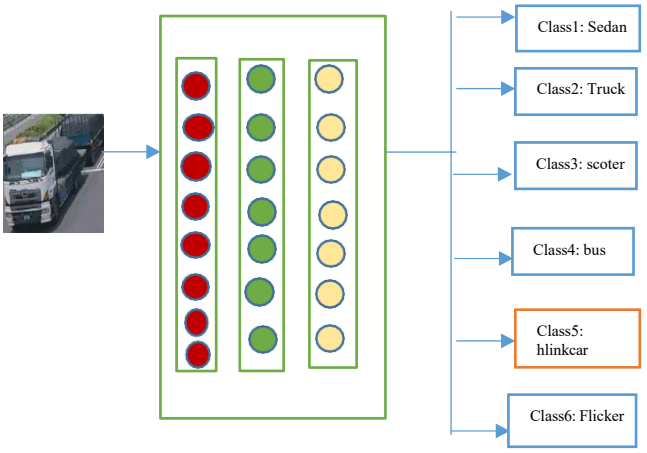


Fig. 3 vehicle categorization is verified using a visual classifier which is mainly working on the YOLO method

7. Vehicle Counting

There are six input movies of roads, each varying in duration from three to five minutes. Moreover, eight films of rainy highways were recorded for testing. This traffic footage was shot at 30 frames per second. Table 2 summarizes first experimental result of vehicle counts using suggested approach in between during day. Small and big vehicles are denoted by the letters S and L, respectively. The proposed approach has a 94 percent vehicle counting accuracy. Table 3summarizes, second experimental result of vehicle counts using suggested approach throughout night. The proposed approach has a 97 percent vehicle counting accuracy. Table 4 summarizes the third trial result of real-time vehicle counting in the rain using the suggested approach. proposed method's vehicle accuracy was 94 percent. picture shows screenshots of a proposed vehicle counting and categorization system with detected cars represent as green rectangular boxes.

TABLE 2: During the day, the suggested approach was usedto count vehicles.

video no	actual n of vehicles			estimated number of vehicles		
	S	L	Total	S	L	Total
1.)	30	6	36	28	5	33
2.)	17	2	19	17	2	19
3.)	15	4	19	15	4	19
4.)	26	3	29	23	1	24
5.)	20	2	22	18	4	22
6.)	18	7	25	19	2	21
7.)	13	3	16	10	1	11

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:02 UTC from IEEE Xplore. Restrictions apply.

TABLE 3: At night, the suggested approach is used tocount vehicles in.

Video No.	actual no of vehicles			estimated number of vehicles		
	S	L	Total	S	L	Total
1)	20	6	26	20	8	31
2)	30	4	34	28	4	32
3)	19	3	20	17	2	19
4)	12	2	14	10	4	14
5)	18	3	21	18	3	21
6)	13	3	16	15	3	18
7)	32	3	35	31	4	35

TABLE 4: On a wet day, the proposed approach was used tocount cars in

Video No	actual no of vehicles			estimated number of vehicles		
	S	L	Total	S	L	Total
1)	9	0	9	9	0	9
2)	12	0	12	10	1	11
3)	13	0	13	13	0	13
4)	7	0	7	8	0	8
5)	11	1	12	14	1	15
6	15	0	15	17	0	17
7	10	0	10	13	0	13
8	7	0	7	10	0	10
9	12	0	12	14	0	14

8. Calculation of speed

The suggested approach is used to estimate the vehicle speed in this subsection. Fig. 5 shows the vehicle's actual and predicted speeds. According to the findings, the absolute percentage average inaccuracy in terms of vehicle speed prediction was around 7.0%. Due to network latency [11], using online video for vehicle speed assessment result in huge speed inaccuracies. As a result, network stability is critical for lowering the percentage inaccuracy in speed estimation.

IV. RESULT & DISCUSSION

The experiments in this paper are performed with the help of Yolo algorithms, OpenCV tool, and the application was created in Python 3 and runs on the Darknet framework. We need to use real-world data to undertake vehicle detection, counting, speed estimation, and other tests to validate a real- time traffic monitoring system and categorization. Three datasets are utilised for assessing suggested strategy in this research.

One of these collections of data was created mostly the video of traffic pictures captured by internet cameras on various Indian roadways. There are 12,000 training photos and 3000 testing[12] images in this set. Second, the Delhi audio and video dataset (DVD) offers information on various level of traffic activity as well as social usage inside Delhi. Finally, there are four categories in the Bengaluru Road traffic monitoring (BAVD) dataset (i.e. car, scooter, truck, big truck, hlinkcar, flinkcar)[14].

The correctness in this article which is mainly working on vehicle categorization in database. If findings of video verification are manual and the recommended system categorization of cars is the same, the count is correct. Otherwise, the car count was incorrect.

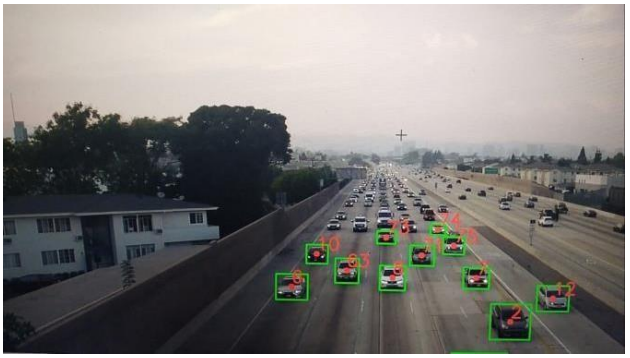


Fig. 4. Final Implemented Software Result

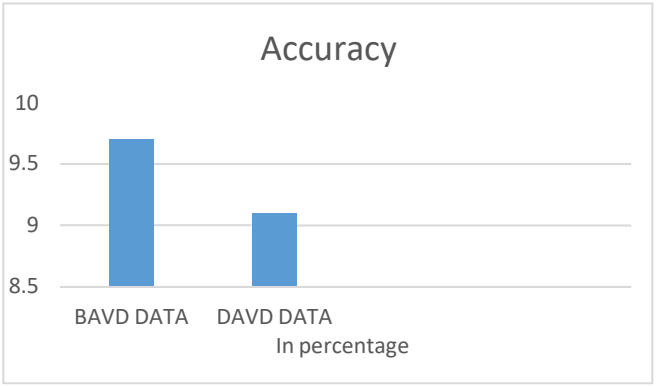


Fig. 5 Comparison of two dataset accuracy

VI.CONCLUSION

For boosting vehicle counts and efficiency, a system which is mainly working on a virtual detection zone [15]. GMM and YOLO is presented in this work. GMM and the virtual zone are utilized to count cars, while YOLO is used to categories them into subgroups. The vehicle's speed is calculating with the help of the vehicle's distance and time travelled. Davd and Bavd datasets are utilized in this work. The YOLO algorithm is used in the suggested technique. On daytime, evening, and rainy days, testing findings show that the suggested approaches have accuracy of 97 percent, 94 percent, and 91 percent, respectively.

There are a few drawbacks to this approach. The automobiles seen in video are considered in a virtual detection zone, which should be large enough to count them. We'll concentrate our efforts in the future on algorithm acceleration and model simplicity.

REFERENCES

[1] Sun, Zehang, George Bebis, and Ronald Miller. "On-road vehicle detection: A review." *IEEE transactions on pattern analysis and machine intelligence* 28.5 (2006): 694-711.

[2] Real-time multiple vehicle detection and tracking from a moving vehicle." *Machine vision and applications 12, no. 2*

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:02 UTC from IEEE Xplore. Restrictions apply.

(2000): 69-83.

[3] Hadi, Raad Ahmed, Ghazali Sulong, and Loay Edwar George. "Vehicle detection and tracking techniques: a concise review." *arXiv preprint arXiv:1410.5894* (2014).

[4] Jazayeri, Amirali, et al. "Vehicle detection and tracking in car video based on motion model." *IEEE Transactions on Intelligent Transportation Systems* 12.2 (2011): 583-595.

[5] Petrovskaya, Anna, and Sebastian Thrun. "Model based vehicle detection and for autonomous urban driving." *Autonomous Robots* 26.2 (2009): 123-139.

[6] Betke, Margrit, Esin Haritaoglu, and Larry S. Davis. "Multiple vehicle detection tracking in hard real-time." *Proceedings of Conference on Intelligent Vehicles*. IEEE.

[7] Bhaskar, Prem Kumar, and Suet-Peng Yong. "Image processing based vehicle detection and tracking method." *2014 International Conference on Computer and Information Sciences (ICCOINS)*. IEEE, 2014.

[8] Sang, Jun, et al. "An improved YOLOv2 for vehicle detection." *Sensors* 18.12 (2018): 4272.

[9] Fang, Jianxin, et al. "A low-cost vehicle detection and classification system based on unmodulated continuous-wave radar." *2007 IEEE Intelligent Transportation Systems Conference*. IEEE, 2007

[10] Kim, HyungJun. "Vehicle detection and speed estimation for automated traffic surveillance systems at nighttime." *Tehnički vjesnik* 26.1 (2019): 87-94.

[11] Chintalacheruvu, Naveen, and Venkatesan Muthukumar. "Video based vehicl and its application in intelligent transportation systems." *Journal of transportation technologies* 2.04 (2012): 305.

[12] Ishida, Shigemi, et al. "SAVeD: Acoustic vehicle detector with speed estimatio of sequential vehicle detection." *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2018.

[13] Wei, Qifan, and Bo Yang. "Adaptable vehicle detection and speed estimation for changeable urban traffic with anisotropic magnetoresistive sensors." *IEEE Sensors Journal* 17.7 (2017): 2021- 2028.

[14] Balamuralidhar, Navaneeth, Sofia Tilon, and Francesco Nex. "MultEYE: monitoring system for real-time vehicle detection, tracking and speed estimation from UAV imagery on edge-computing platforms." *Remote Sensing* 13.4 (2021): 573.

[15] Zhou, Jie, Dashan Gao, and David Zhang. "Moving vehicle detection for automatic traffic monitoring." *IEEE transactions on vehicular technology* 56.1 (2007): 51-59.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:02 UTC from IEEE Xplore. Restrictions apply.

DESIGN AND DEVELOPMENT OF AN AUTONOMOUS UNDERWATER VEHICLE - LAPRAS 2.0

Sriram Chava, Saurav Ranjan, Shivang Dudani, Utkarsh Verma, Shorya Kumar, Kritika Bansal, Rajat, Sundeep Chand, S. Suchit, Chirag Garg, Suryansh Gangwar, Ishpreet Singh Luthra, Shivam Rastogi, Abhinav Agrawal, Rakshit Garg, Dhruv Upreti, Aaryaveer Bhardwaj, Aaryan Mendiratta, Dev Verma, Harjas Singh, Vidushi Saxena, Hriday Bhardwaj, Parth Sharma, Anupam Maurya, Shubh Mittal

Delhi Technological University, Shahbad Daulatpur, New Delhi-110042, India

ABSTRACT:
LAPRAS is the flagship Autonomous Underwater Vehicle developed at Delhi Technological University.Using the insights from Robosub 2021 , Lapras 2.0 has enhanced upon the capabilities of its predecessor. Manipulators have been upgraded for better dynamics in autonomous operations.Vision algorithms have been finely tuned which has led to a massive improvement in the success rate of object detection .

I. COMPETITION STRATEGY

We chose the G-Man side to complete our tasks as the distinct colour schematic of the G-Man side made it a source to provide more accurate vision algorithm results. For the gate task, when the AUV is in yaw manoeuvre trying to find the direction to surge into to cross the gate, we simply detected the orange ends of the gate and calculated the relative difference between both the sides to transitionally and rotationally align our bot and move towards the centre of the gate. When the distinct features of Bootlegger and G-Man come into the view of the camera, the algorithm shifts to the detection using the models we trained for object detection in YOLOv5. For the path marker, after identifying it using colour thresholding, we applied the concept of HoughLinesP and HoughCircles to detect all the 3 edges of the path marker and give a path to the bot.

For the buoy task, the correct side is identified using model training we implemented using YOLOv5 and then the distance estimation algorithms ensure that the bot touches the buoy and returns back.

For the bins task, firstly the bot searches for the location of the bins using the colour of the lid, then the correct side is identified using the object detection and then the marker is released, the bottom camera and colour thresholding algorithms are used to check whether the marker was dropped in the correctly.For the torpedo task, the correct side is identified using object detection and using contour detection and the comparison of areas, we identify the smaller and bigger openings.

For the octagon task, first the AUV detects the table and then the bottles which it later picks up by the gripper, after surfacing, AUV performs the yaw manoeuvre to find the correct table to place the bottle.

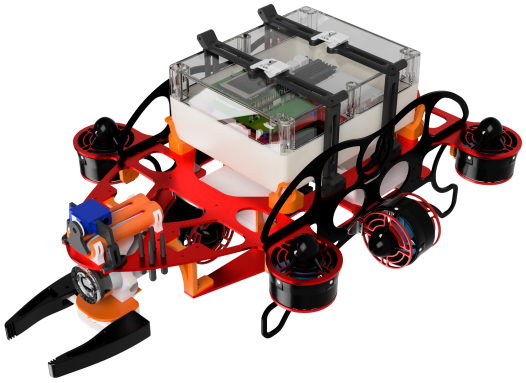


Figure 1.1 - Cad render of Lapras 2.0

II. DESIGN RATIONALE

A. MECHANICAL DEPARTMENT

1) Frame

The frame is the primary structure in a vehicle, providing support to all mounted entities. The overall frame features 4 parts: 1 base plate, 1 battery support plate and 2 side plates. The base and battery support plates are connected through custom 3D printed ABS mounts. The base plate holds the main hull, front camera hull, 4 heave thrusters, 2 sway thrusters, torpedo launcher, front camera and gripper assembly. Each side plate holds one surge thruster each. The battery support plate holds the battery enclosure, bottom camera hull and dropper assembly. Eight thrusters are used to get control over all the six degrees of freedom of the AUV.

2) Manipulators

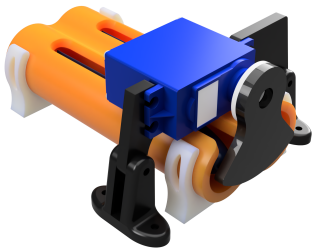


Figure 2.1: Design Render of Torpedo Launcher

The design of Lapras 2.0 torpedo launcher is very compact due to incorporation of a spring-launched mechanism for providing launch thrust. The assembly has 2-parallel tubes for holding the projectiles placed adjacent to each other. A servo is mounted above the tube and is fitted with an Aluminium 6061-T6 obstruction plate. The servo rotates the obstruction plate by 30 degrees in either of two directions, clearing the opening of a barrel and the torpedo is launched by compressed spring.

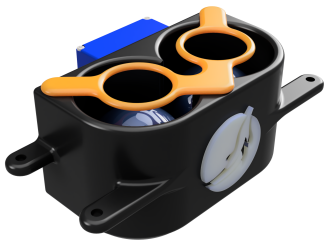


Figure 2.2: Design Render of Dropper

The dropper assembly is powered by a single servo mounted behind the marker housing. Both dropper tubes placed adjacent to each other ensuring high accuracy during dropping sequence. A marker is dropped by 30 degrees rotation of a heart-shaped obstruction plate in either of two directions, clearing the path for its drop. The dropper body is made up of 3D printed ABS plastic.

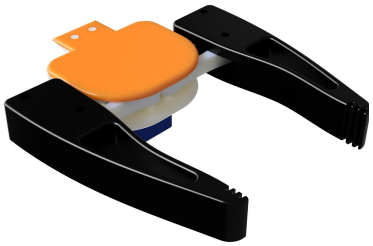


Figure 2.3: Design Render of Gripper

The design of the gripper focuses on simplicity, reliability and modularity. The outcome of the design process is a 2 claw parallel gripper. It features 2 claws, 2 driver links and 2 coupler links. The driver links have spur gears at the ends and are meshed together at the base, the coupler links do not possess any gear and their primary function is to support the claw and restrain its movement. The design allows for gripping and object upto 90 mm wide. The entire assembly is 3D printed using ABS plastic. Compared to our previous rotatory fingers design, this parallel gripper is capable of gripping larger and more complex shaped objects.

B. ELECTRONICS DEPARTMENT

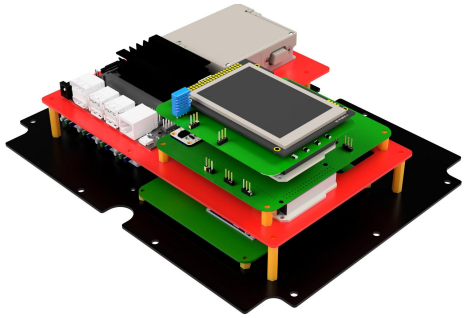


Figure 2.4:Electronics Stack

The embedded stack has been developed keeping in mind design rationales such as robustness,

modularity, space constraints and ease of debugging.

1. Power

A Power Distribution Board (PDB) was designed to power all the electronics inside the vehicle. The 12VDC input is used to power the GPIO board. PDB has buck converters that convert the input voltage to 7V and 5.1V respectively. A DPDT switch has been used as a kill switch, connected between the GPIO board and the PDB, to halt the operation of the vehicle.

2. Battery monitoring system

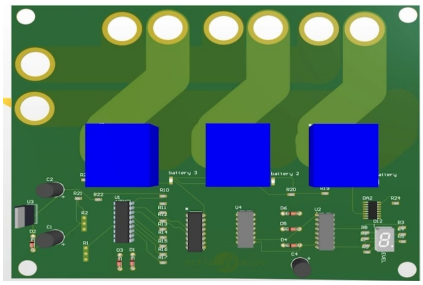


Fig 2.5 Battery Monitoring System

Three Lithium polymer batteries (4 cells) with a total capacity of 30000mAH were chosen for the input power. A custom battery monitoring system PCB was designed, which continuously monitors the cell voltage and the total voltage of each battery and cuts it off from the circuit as soon as it reaches the threshold voltage (3.1V). The system is low power, cheap, and provides a visual indication of the battery levels and the battery currently in use.

3. GPIO board

A GPIO board has been developed for the integration of sensors and actuators. The Arduino Mega 2560 has been used as the microcontroller and is powered through the PDB. In turn, the Arduino powers the sensors on the GPIO board, receives the sensor data and provides the PWM inputs to the Electronic Speed Controllers and the Servo Motors as well.

For the sensors, we have used ACS712 Current Sensor, Bosch BNO055 IMU, and the BAR 30 pressure sensor. Here, the current sensor has been used to monitor the current draw of the servo motor, using which we can detect whether an object has been gripped or not.

Appropriate connections have been made to provide the required inputs for the debug board. Apart from this, the GPIO board also powers the ESCs through the PDB to ensure easier and more flexible connections.

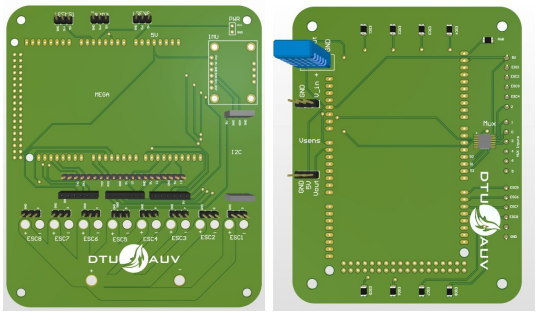


Fig 2.6 GPIO and Debug board

4. Debug Board

A Debug Board was developed to make our system more interactive with the use of an LCD Screen and several LEDs. The Debug board receives several inputs from the GPIO board including PWM inputs to the ESCs, the voltage across several components on the GPIO board, and that of the DHT11 temperature sensor on the Debug board. These are checked to ensure appropriate operation of the AUV An 8-to-1 Multiplexer IC has been used to interface these with the voltage sensor. This ensures fewer components and hence a compact system. Several LEDs have also been used to indicate the ESCs operating at any given time.

5. Acoustics

Fig 2.7 Hydrophones Placement

The hydrophones have been closely placed in a square array to avoid spatial aliasing. A custom PCB consisting of an amplification unit and a bandpass filter is used for signal conditioning before sending the signal to the DAQ.We have used the NI9223 DAQ as our analog input module for data acquisition .The signal will then be used to

find the direction of arrival using the MUSIC (Multiple Signal Classification) algorithm on the single board computer.

C. SOFTWARE DEPARTMENT

1. Computer Vision Algorithms

To complete the Robosub challenges, we made sure that the vision algorithms put in place were efficient, feasible in the real environment and yet were highly accurate. We differentiated the props from the surrounding environment on the basis of their colour and then after putting a filter on the basis of area, extracted out the contours of our target. Further on, bounding rectangles helped us to navigate the bot as needed accordingly. In order to navigate in between two tasks, we implied certain algorithms to navigate along the path marker. The path marker prop was divided into two subcategories - 1) The semi circle ends, 2) The linear edges. Detection of both these subcategories were handled differently. The technique implied was Hough transform which can be used to detect any shape, if you can represent that shape in mathematical form. Assuming the intersection point as a vertex and the centres of the semicircular ends as other two vertices , an isosceles triangle was constructed. The slopes of the detected lines were used to find the internal angles of this triangle and then by using this along with the distance between the two centres, through simple trigonometric relations, we could find the coordinates of the intersection point which helped us to draw a line along the entirety of the path so as to easily navigate.

Figure 2.8 Detection of path marker and calculating the angle

2. Detecting the Targets

This year's iteration brought new innovations in computer vision and deep learning algorithms, the most significant one being SRGAN. Super-Resolution Generative Adversarial Network (SRGAN) architecture allows us to artificially enhance image resolution with minimal computations. Utilising this technique, we were able to enhance our camera feed by up to 4x its current pixel space, allowing better execution of other CV algorithms. We have also utilised state-of-the-art object detection algorithms such as YOLO-V5 and fasterRCNN to detect and classify different tasks corresponding to G-man and Bootlegger.

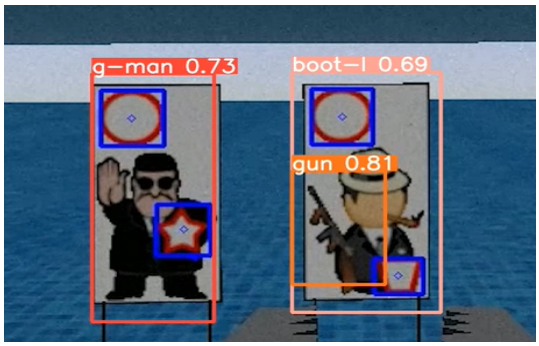


Figure 2.9 Identifying and classifying the targets in torpedo task

3. Simulations

Simulations are an essential tool to determine the effectiveness of our vehicle before testing it in a physical environment. For simulating the vehicle, we used the UUV package in Gazebo simulator. Adding different types of noise to the environment helped improve accuracy of the simulation.

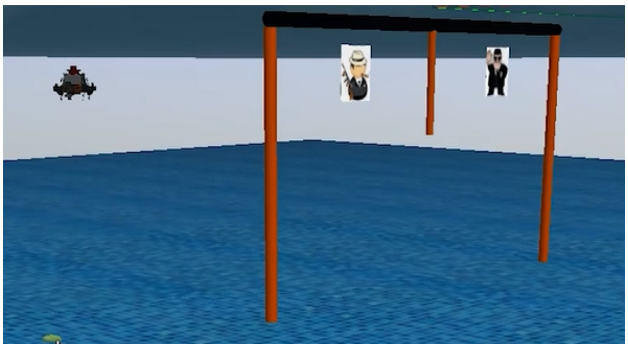


Figure 2.10 Gazebo scene setup for the gate task

4. Mission Planner

We used SMACH as a tool for our mission planner. SMACH gives a full introspection into the state machines, state transitions, and data flow. SMACH is a library for task-level execution and coordination, and provides several types of "state containers". The mission planner helps to make decisions based on the input given by the camera feed filtered by the vision algorithm and plan its journey accordingly based on the hierarchical state-provided.

III. EXPERIMENTAL RESULTS

As we participated as an online participant, we tested the working, control, accuracy and efficiency of our bot using the Gazebo simulation environment. These simulations proved to be immensely useful in providing some insights that we would take care of during the production of our future bots. We have implemented Computer Vision as well as Machine Learning algorithms using OpenCV and YOLOv5. The use of both these technologies proved a very significant improvement from our last year's bot.

So, in order to compete with a near equal level to in person teams, we needed to simulate our bot, its environment and its working in as realistic environment as possible. The simulations needed to be built in a robust manner along with incorporating some real life features like criss cross design at the bottom of the pool, random waves, dirty water and random noise. These features also made it possible for us to predict the accuracy and success levels of various tasks that our bot can perform.

1) For the gate task along with style, success rate would be near to 99%,
2) As we have performed the path marker and buoy tasks in physical mode as well created their simulations, their accuracy level would be 100%,
3) As we have incorporated model training in YOLOv5 as an improvement to last year's feature matching algorithm for object detection, it has increased the accuracy levels of buoy and torpedo tasks by an approximate of 15% each, making the success rate of buoy task to be 95%, shooting the torpedo at the bigger opening to be 85% and shooting the torpedo at the smaller opening to be 85%.
4) Significant improvement in the design of the gripper with respect to the thickening of its fingers and addition of groves at the tips, made it easier to lift the bottles, moreover the incorporation of Machine Learning technologies for detection of correct tables increased the octagon task's accuracy level to 65%.

The water tight hulls used in AUV already possess IP-68 rating and were tested for any manufacturing defects. Holes were drilled in the lower half of the hull using 10mm drill bits and were fitted with potted cable penetrators by BlueRobotics. Nitrile rubber O-rings were placed in the penetrators' grooves because of their ability to expand under differential pressure and completely fill the grooves. The hull was then completely closed from top and held underwater for long durations to check for any leakage. Since we are participating as an online contestant, the gripper was simulated to hold objects and task specific props of varying dimensions. Then the gripper was simulated under varying loading conditions such as static, dynamic and impact loading to optimise the weight to strength ratio.

IV. ACKNOWLEDGEMENTS

DTU AUV would like to thank and express gratitude towards everyone who has come together and helped us to make the project a huge success. We'd like to express our sincerest thanks to our faculty advisor Dr Ajeet Kumar and his guidance on the design considerations. Moreover, the team wouldn't have been able to function without its cold sponsors who generously funded us and helped us acquire all the resources that were required during

the many stages of the project. Finally, we would like to thank the knowledgeable and resourceful network of DTU AUV alumni who made themselves available for advising and sharing their valuable experiences and helping us circumvent multiple roadblocks that one faces while designing a vehicle.

V. REFERENCES

[1] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice. Initial results in underwater single image dehazing. In Proc. MTS/IEEE OCEANS, 2010.

[2] Natu, Aditya, Vipul Garg, Purnyatre Gaur, Piyush Raj, Dhruv, Shivam Sain, Upasana Biswas, Deepanshu Bansal et al. "Design and Development of an Autonomous Underwater Vehicle VARUNA 2.0."

[3] Wang, G., Zhou, J.Q., Zhao, Y.T., Feng, J. and Sun, Q., 2021. Performance Degradation Modeling and Reliability Analysis of Aviation Connectors in Marine Environment. In IOP Conference Series: Materials Science and Engineering (Vol. 1043, No. 3, p. 032029). IOP Publishing.

[4] M. M. M. Manh˜aes and S. A. Scherer and M. Voss and L. R. Douat and T. Rauschenbach, UUV Simulator: A Gazebo-based package.

APPENDIX A: COMPONENT SPECIFICATIONS

Component	Vendor	Model/Type	Specs	Custom/Purchased	Cost	Year Of Purchase
Buoyancy Control	NA	NA	NA	NA	NA	NA
Frame	In House	Plane truss structure 3D printed attachments	4mm Aluminium 6061 T6 sheets 3D Printed ABS	Custom	$150 $81.20	2020
Waterproof Housing	Hammond Hammond Polycase	1554YA2GYSL 1554C2GYCL AN-07F-01	IP68 rated, Polycarbonate IP68 rated, Polycarbonate IP68 rated, Aluminium	Purchased (Modified)	$108 $27 $33.69	2020
Waterproof Connectors	Blue robotics	Potted Cable penetrator	Quantity: 20,M-10 thread, 6mm cable	Purchased	$100	2019
Thrusters	Blue robotics	T200	Quantity: 8,Operating Voltage:7-20V, Full Throttle Current(@20V):32A	Purchased	$1432	2019
Motor Control	Blue Robotics	ESC for T200	Quantity:8,7-26 volts (2-6S)	Purchased	$216	2019
High-Level Control	Arduino	Mega 2560	6-20V,54 digital I/0 pins,16 analog pins, Clock speed:16MHz	Purchased	$24.30	2019
Actuators	Hitec	Servo HS5646WP	IP67, Torque:11.3-12.9 Kg-cm^2/s^2	Sponsored	Sponsored	2019
Battery	Tattu	Tattu 14.8V 25C 4S 10000mAh Lipo Battery Pack	14.8V 25C 4S 10000mAh	Purchased	$350	2019
Converter	Traco	THM 60 – 2411WI	Output Voltage:5.1, Current:12A	Purchased	$158	2021
Regulator	Robu	LM2596S	Quantity:2,Input Voltage:3-40V,Output Voltage:1.5-35V(Adjustable)	Purchased	$1.60	2020
CPU	Nvidia	Jetson Nano	Quad-core ARM Cortex-A57 MPCore processor	Purchased	$300	2020
Internal Comm Network	N/A	N/A	N/A	N/A	N/A	N/A

External Comm Interface	N/A	N/A	N/A	N/A	N/A	N/A
Inertial Measurement Unit (IMU)	Bosch	BNO055	9 Axis, supply voltage:2.4-3.6V, Range:125-2000 deg/s	Purchased	$10.70	2019
Manipulator	In House Developed	Gripper Dropper Torpedo Launcher	3D Printed ABS	Custom	$24.33 $13.10 $15.73	2020
Algorithms: vision	N/A	OpenCV, YOLOv3	Underwater Single Image Dehazing, YOLO Object Detection, Distance Estimation using triangulation, ORB	Custom	N/A	N/A
Algorithms: acoustics	N/A	TDOA	Hyperbolic positioning	Custom	N/A	N/A
Algorithms: localization and mapping	N/A	OpenCV	Localisation using vision algorithms	Custom	N/A	N/A
Algorithms: autonomy	N/A	Smach	Finite State Machines	Custom	N/A	N/A
Vision	Logitech	C930E	Full 1080p HD video at 30 frames per second	Purchased	$ 130	2019
Acoustics	Teledyne	Reson TC-4013	Quantity:3,1 Hz-180KHz ,depth:700m	Purchased	$538x4	2020
Algorithms: vision	N/A	OpenCV, YOLOv3	Underwater Single Image Dehazing, YOLO Object Detection, Distance Estimation using triangulation, ORB	Custom	N/A	N/A
Algorithms: acoustics	N/A	TDOA	MUSIC	Custom	N/A	N/A
Algorithms: localization and mapping	N/A	OpenCV	Localisation using vision algorithms	Custom	N/A	N/A
Algorithms: autonomy	N/A	Smach	Finite State Machines	Custom	N/A	N/A
Open-source software	OpenCV, YOLOv3, Robot Operating System, Smach, Gazebo, Python, C++					

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

Disease Detection in Maize Plant using Deep Convolutional Neural Network

Sohum Gupta
Applied Physics Department
Delhi Technological University
Delhi, India
26sohumgupta@gmail.com

Josh Vishnoi
Applied Physics Department
Delhi Technological University
Delhi, India
joshvishnoi@gmail.com

A.S. Rao
Applied Physics Department
Delhi Technological University
Delhi, India
drsrallam@dce.ac.in

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835737

***Abstract*— Maize is one of the most versatile crops. It has various types, used in many industrial processes and has wider adaptability under varied agro-climatic conditions. In 2020. a total of 145+ million metric tons of maize was produced by 160+ countries making it the second-most widely cultivated crop in the world.**

In 2020-21, India cultivated 30 million tonnes of maize. To meet this high demand, precision farming techniques are used for identifying, analysing and managing crop quality, helping generate optimum production as well as saving farming resources.

One of the most significant reasons that affects crop output in India is the late detection of diseases in maize plants. This research will help in timely detection of such diseases and assist the farmers in the cultivation process, thus directly impacting maize production output. Maize plant leaf diseases like Leaf Blight. Leaf Spots, and Leaf Rust have clear cut symptoms that can be visually identified. Thus such diseases can be identified and their classification can be done using Artificial Intelligence. In image classification and pattern recognition, Deep Convolutional Neural Networks are known to have higher efficiency and accuracy. Hence it can be used as a potent tool in the diagnosis of maize plant diseases. Thus, this research work has proposed a Deep Convolution Neural Network model which combines with Transfer Learning technique, resulting in reduced training time taken by the AI model. For higher accuracy acquisition, Data Augmentation techniques are used to improve upon the variability in the dataset that is being used to train the network. This model that has been put forward achieved an accomplished accuracy of 97.93%.

Attaining a greater accuracy is really vital for the amelioration of present crop wastage conditions due to diseases that are left undetected. This accuracy can even be increased by adding some other parameters about which will be discussed in future works section.

Keywords— Deep Convolutional Neural Network, Data Augmentation, Maize Plant Disease, Transfer Learning.

I. INTRODUCTION

Maize, also known as the 'Queen of Cereals' is cultivated in about 1.5m km^2 in about 160+ countries and contributes around 36% in total cereal production around the world. The United States of America is the largest producer of Maize (35%) in the whole world followed by China. It is also its largest consumer and exporter.

In India, maize also known as Indian corn, is the third most important food crop after rice and wheat as it covers approximately 80% area and its production is around 27.8 million mega tonnes during the year 2018-19[1]. India also ranks fourth in area and seventh in production of maize in the world which is around 2% of world's production. maize also contributes around 9 % to the national food basket. In addition to quality feed for animals and staple food for human beings, maize also serves as a basic raw material as an ingredient to hundreds of industrial products that includes alcoholic beverages, pharmaceutical products, different types of oils, artificial sweeteners, cosmetic items, textiles and paper industries etc.

Approximately 47% of maize produced is used as poultry feed. Of the rest of the produce, 13% is used for livestock feed and food purposes, 12% for industrial purposes, 14% in starch industry, 7% as processed food, and 6% for export and other purposes. There are more than 15 million farmers in India cultivating maize and generating employment for more than 650 million people[2].

When compared to world average, Indian Maize performs comparatively low, in spite of the fact that a significant population is dependent on it for employment and usage.

Maize plants are highly susceptible to a number of diseases and pests. Worldwide, there is a loss of around 20-40% of crops due to leaf diseases which costs $220 billion to the global economy[3]. Pathogenesis stage of the disease is the most important factor in the success of chemical control. Ignorant and careless use of pesticides may increase the resistance to germs over time, severely disrupting the plants' ability to fight. Therefore, accurate and timely diagnosis of plant diseases is an important factor in precision farming.

Maize leaf when attacked by such biotic stresses can be identified by the use of soft computing which plays the most crucial role. New techniques are being continuously developed to identify such diseases, such as machine learning, deep learning, computer vision, image processing, generative modelling etc[7]. Due to their low efficiency,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:50 UTC from IEEE Xplore. Restrictions apply.

these models are not being adopted at large scale. There are many leaf and plant diseases which cannot be identified visually but most common leaf diseases can be identified visually. For example, leaf rust produces dark brown coloured rust, elongated spots on both leaf surfaces. The spots contain rust spores that are coloured cinnamon brown, and form a big round blot that turns from orangish to reddish to brownish with a yellowish border with time[10].

Classification of images for the past few years has been completely governed by the field of Deep Learning, thus the proposed model helps to determine maize plant diseases and classify them into four categorical classes – Healthy Leaf, Diseased/Unhealthy Leaf with diseases such as Leaf Spots, Leaf Blight or Leaf Rust[8].

The further paper is organized as per the following order: Different architectural components working and methodologies of features used in our model are explained in detail in Section II. Dataset used in this research work is explained in Section III in detail. The experiments performed by us are described in Section IV. Final results of our experiments are presented in Section V. Conclusion along with future work is presented in Section VI.

II. WORKING AND METHODOLOGIES

The deep learning model, Convolution Neural Network or CNN is used which is a variant of neural network that is mostly used in problems related to Computer Vision. It finds its usage in most of the image classification problems. The hidden layers of a convolution neural network consist of convolution layers, fully connected layers, pooling layers and other normalizing layers.

Our proposed model is implemented in the python programming language with TensorFlow Keras library which uses sequential classification. This deep learning model makes use of Mini Batch Gradient Descent Optimization to modify and update approximately 19 million parameters. We used general gradient descent algorithms in the initial phases of development and found mini batch optimization to be highly performant compared to them. To train the neural network, the images are fed into the model in batches of size 32. The epoch number is kept at 22. Fig 1 shows the proposed network's architecture in detail. The pictures are classified into 4 categories and a total of 1488 images with 256x256 pixel dimensions are used. In the subsequent sections, transfer learning and data augmentation are explained. Proposed model utilises these techniques to achieve a high success rate.

A. Transfer Learning

It is a machine learning style where we develop a model for a task then we further use that model as a foundation for another model. Such kind of pretrained models are also used in the field of deep learning for applications like NLP (natural language processing) or computer vision applications. Example: As we know that both piano as well as violin are musical instruments, if a person who has expertise in piano can very easily learn violin because of his knowledge of playing piano and vice versa. Learning to transfer, according to psychologist C.H.Judd's generalisation theory of transfer, is the outcome of generalisation of experience. According to this hypothesis, there must be a link between two learning processes for transfer to occur. Knowledge transfer might be unsuccessful if domains have little in common. We have used the InceptionV3 model as a starting point for our model[12].

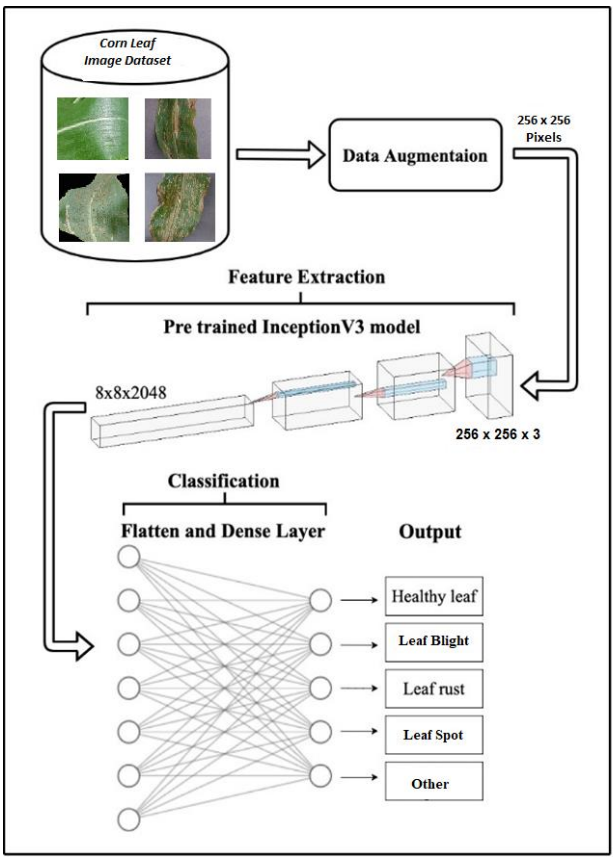


Fig. 1. Network Architecture

Deep Neural Network consisting of 48 layers has been used. More than a million pictures from round about 1000 different classes from the ImageNet Database have been used to train this model. It comprises around 25 M parameters. For our model, Google's TensorFlow library is used for transfer learning. For the classification part, a flatten and a dense layer is added to our model, which varies the number of parameters of our model and puts a stop to the issue of overfitting as well. Furthermore, Cross-domain resemblance may most certainly not necessarily aid our learning process since cross-domain similarities might be deceptive. Apart from this, our concept of prior experience having an opposite influence on learning new tasks is called

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:50 UTC from IEEE Xplore. Restrictions apply.

negative transfer in psychology. Transfer learning may be further separated into two groups based on the disparity across domains: homogeneous and heterogeneous transfer learning[5][9].

B. Data Augmentation

Data Augmentation is a method to increase the amount of data by modifying the existing data or creating new data from the existing data. The process consists of using varied augmentation methods such as rotating, scaling the image, changing the hue, brightness of the image, filtering, blurring the image to produce modified versions of the original image without changing their labels[6]. This ensures that the model is trained on a huge variety of enriched data and picks up more feature learning in a sturdy fashion. It puts a stop to the issue of overfitting as well.

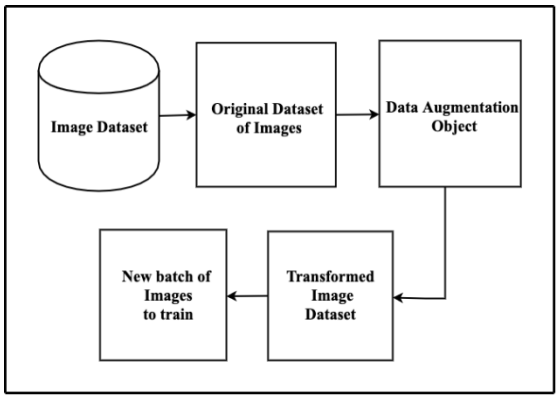


Fig. 2. Augmentation techniques on the set of images

This method described above can only help in enlarging the dataset but it cannot be used to replicate the diversity found in nature. So, instead of considering the entire leaf only the image containing symptomatic parts is used. In the proposed model, the transformation is applied at the runtime ensuring that a new transformed image is input at every epoch into the model, which increases the accuracy of our model. This technique is generally called in-place data augmentation[4].

C. Multiclass Categorical Cross-Entropy

To evaluate the performance of any machine learning model, a loss function is used. The loss function results in a larger value if the predictions from the model are greatly deviated from the actual values.

So, we need to minimize the value of loss function to improve our prediction accuracy. To minimize the loss, we can use one of the many popular loss functions available such as Mean Absolute Error, Mean Squared Error, Categorical Cross Entropy, Cosine Similarity, Binary Cross-entropy etc. One of the most used loss functions for multi-class classification is the Categorical Cross Entropy Loss Function. This function compares the predicted value from our model with the actual values. This function is based on the assumption that only one class is correct out of all possible ones so the predicted class is taken as 1 and all others are kept 0. Its equation is given as:

$$L(y,\hat{y}) = -\sum_{j=0}^{m} \sum_{i=0}^{n} (y_{ij} * log(\hat{y}_{ij})) \qquad (1)$$

where the predicted value is represented by $\hat{y}$, and actual value is represented by y; n represents the total number of classes and m represents the total number of data points.

D. Optimization Algorithm

To minimize a given mathematical function g(x), we need some Optimization Algorithm. g(x) usually depends on some values, which in deep learning models are its parameters and optimization algorithms assist in calculating the output efficiently[17]. Gradient Descent is one such popular first-order optimization. It is given as:.

$$\theta = \theta - \alpha \nabla_{\theta} C(\theta; x_o, y_o) \qquad (2)$$

where (x_o, y_o) constitutes the points from the learning set and the optimization parameter is represented by θ; C (θ; x_i, y_i) is cost function.

Our dataset has been split into batches of 32 pictures each and the model uses a Gradient Descent algorithm known as Mini-Batch Gradient Descent Algorithm. One major advantage of using a mini batch algorithm is that it consists of advantages of the two algorithms that are Batch Gradient Descent and Stochastic Gradient Descent algorithms. In the Batch Gradient Descent algorithm, taking one step towards the direction of minimum, it runs on complete data which is a disadvantage, but with Stochastic Gradient Descent, the model calculates the gradient of each randomly selected data point which result in high speed but loses the positives of vectorization. It might also be error prone in some cases.

III. DATASET

The dataset [11] has a diverse set of maize leaf images for effective diagnosis of diseases. The dataset is pre-processed to increase the diversity and get rid of undesired details which enhances leaf's symptomatic features.

The set of images consists of diseased and disease-less leaves that are afflicted by various forms of stressors. The set of images is bifurcated into the symptom dataset containing pictures of leaves with symptoms with the diseased part being zoomed in the images and the healthy leaf dataset which has the unedited images.

Leaf dataset: It consists of a total 1747 maize leaf images along with a CSV file for different biotic stresses labels. The unedited leaf images are of the dimensions - 256x256 picture elements.

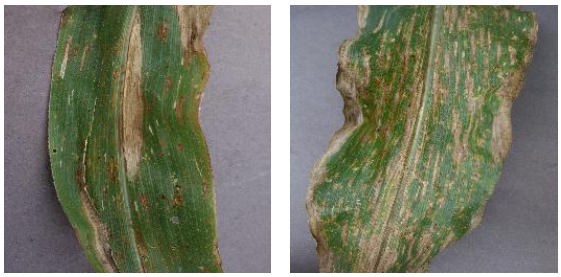
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:50 UTC from IEEE Xplore. Restrictions apply.

Fig. 3.Illustrative set of healthy maize leaves images

The leaf images are divided into four different categories - Healthy Maize leaves, Infected Maize leaves with Leaf Spots, Leaf Blight or Leaf Rust. .

Symptom dataset: This dataset consists of a total 2147 images which are cropped to contain the symptomatic parts of leaves present. The set of images with symptoms has been divided on the basis of leaf diseases - healthy leaves, Leaf Spots, Leaf Blight, Leaf Rust. The above four categories are used for categorization of dataset images. This collection of images are divided into further parts for our model, i.e. training, validation and test folders.

(a) Disease Free leaf (b) Rust



(c) Leaf Blight (d) Spots on leaf

Fig. 4. Some illustrative set of pictures of leaves with symptoms with the diseased part being zoomed in

IV. EXPERIMENTS

For implementing our model, the dataset is divided into three parts, that are training set, validation set of images and testing set. Of all the images 70% have been taken up for training, one half of the remaining for testing and the other half for validation purposes[13]. We have also used data augmentation to diversify our dataset. The output images are of dimension 256 X 256 pixels. Pre-trained InceptionV3 model is used, as it increases the efficiency of our model and reduces the training time[14]. The convolutional layers used in it have weights and biases pre-trained. New images produced after each epoch are passed to the pre-trained InceptionV3 model.

In our sequential model, we have also added a dense and flatten layer which helps in classification. The multidimensional output from the convolutional layer is converted into a linear output by the help of a flattening layer. It is done because its output can be directly fed to a dense layer. Then, the calculations are performed by the dense layer and the Softmax Activation function present in it gives the output[15].

Loss function that is being used in our model is a multi-class loss function known as Categorical Cross Entropy. The optimization algorithm being used in our model is a mini batch gradient descent algorithm. Batch size (that is the number of training examples used in one iteration) of 32 has been used for obtaining maximum accuracy, as small batch size generally results in reduced accuracy.

The model is also tested across different values of Step Size(or learning rates) to get the best accuracy. The results of testing different learning rates are mentioned below in table 1:

TABLE I. LEARNING RATE VERSUS ACCURACY & LOSS VALUE

Step Size	Loss Value	Accuracy(%)
0.0001	0.47	95.84
0.001	0.45	95.31
0.005	0.36	97.24
0.01	0.34	97.93
0.1	1.68	73.58
1	10.1	38.10

From our experiments, we found that with an increase in learning rate/Step size, our model's accuracy increased, reaching a maximum of 97.93%, and then started to decrease further. So, the Step Size of our model is concluded to be 0.01.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:50 UTC from IEEE Xplore. Restrictions apply.

V. RESULT AND DISCUSSION

An overall accuracy of 97.93% is achieved using the proposed deep learning model, with a loss value of 0.34.

For each successive iteration, we found that the accuracy of the model increases and the loss value decreases. Moreover, the accuracy value increases and loss decreases significantly with every epoch. Table II shows the changes in accuracy and loss value of the model with successive iterations. We notice that after the 10th iteration, no major improvements in the accuracy of the model is observed even though the loss values decrease minutely.

TABLE II. EPOCH VERSUS ACCURACY & LOSS VALUE

Epoch value	Loss Value	Accuracy (%)
1	1.59	63.58
3	0.73	72.93
6	0.39	93.57
8	0.37	96.98
10	0.34	97.93
13	0.34	97.93

As can be seen in the figure below the value of loss value and accuracy not varying much after the epoch value of 6.

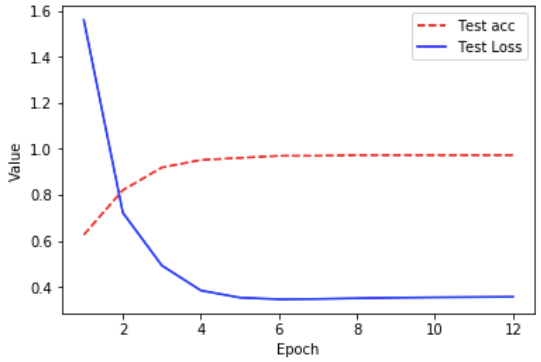


Fig. 5. Epoch v/s Accuracy and Loss Values graphical representation

To avoid the problem of overfitting of data, we have not used maize plant leaf dataset directly. The technique of data augmentation to generate a more varied training dataset is used which resulted in a much more accurate and reliable model.

In order to improve on the training time and make it more faster and efficient, the model uses Transfer Learning method. To improve the classification process of our neural network model we have also included a flattened layer and a dense layer in the model.

Final values of the model hyper parameters for the proposed neural network are presented below in Table III.

TABLE III. MODEL HYPER PARAMETERS VALUES

Model Parameters	Values
Activation function	Softmax Activation Function
Loss function	Multiclass cross entropy
Optimization Algorithm	Mini batch gradient descent
Step Size(or Learning rate)	0.01
Batch size	32
Momentum	0.9

Some Limitations of our system

1) If any image being used has considerable noise then our model sometimes identifies the image as a totally dissimilar image.

2) If a leaf contains multiple visible diseases, and it is passed in our model, then it can not be said for sure that our model will detect all the diseases but will definitely detect one disease.

3) Training time for our model is quite high especially due to CNN requiring big datasets. One would require some devoted hardware (e.g. GPU, TPU) for reducing it.

VI. CONCLUSION AND FUTURE WORK

Usage of convolutional neural networks to detect and categorise healthy and unhealthy maize plant leaves is demonstrated in the research paper. The diseases are categorised into five major classes – Healthy Leaves, Diseased Leaves, Diseased leaves further divided into Leaf Spot, Leaf Blight, Leaf Rust, and other diseases, which are among the most widespread diseases in maize plant leaves. Proposal of deep convolutional neural network for early detection of diseases in maize plant leaves efficiently and with a high level of accuracy is one of the major contributions of our research work. This detection model owns its importance because of the huge amount of crops that are wasted due to certain diseases that are left undetected.

To summarize -

1. The convolutional neural network proposed achieved a very accurate prediction of 97.93% in early detection of most widespread maize plant leaf diseases.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:50 UTC from IEEE Xplore. Restrictions apply.

2. Maize plant farmers and cultivators in India can be supported with the early detection of its disease using the proposed approach which in turn will increase the country's overall maize production.

To further improve the prediction rate and achieve an even higher success percentage for the diseases in maize plants, the model will require additional input data such as humidity, air temperature, location, plantation time and weather conditions.

It's really crucial for both the farmers individually and the whole country that these production problems are addressed more so that present conditions are ameliorated. In the near future, we are aiming to incorporate these requirements and make necessary revisions to the proposed model in order to take various atmospheric factors into account and also factor in various other new and prevalent diseases.

ACKNOWLEDGEMENT

The authors are indebted to the central library of Delhi Technological University for providing all the research materials and lab access for smooth and successful functioning of the project and research work, besides it, we are sincerely thankful to our supervisor prof. A.S. Rao for mentoring and guiding us throughout the project and research work.

REFERENCES

[1] C.M. Parihar, S. L. Jat, A.K. Singh, R. Sai Kumar, K.S. Hooda, Chikkappa G.K. and D.K. Singh. Maize Production Technologies in India. DMR Technical Bulletin2011/---. Directorate of Maize Research, Pusa Campus, New Delhi-110 012. Pp 30.

[2] D. P. Chaudhary, Ashwani Kumar*, Sapna S. Mandhania, P. Srivastava and R. S. Kumar. Maize As Fodder? An alternative approach, Directorate of Maize Research, Pusa Campus, New Delhi - 110 012, Technical Bulletin 2012/04 pp. 32. * Scientist, IGFRI, Jhansi.

[3] Khan, Masudulla, Azhar U. Khan, Mohd A. Hasan, Krishna K. Yadav, Marina M.C. Pinto, Nazia Malik, Virendra K. Yadav, Afzal H. Khan, Saiful Islam, and Gulshan K. Sharma. 2021. "Agro-Nanotechnology as an Emerging Field: A Novel Sustainable Approach for Improving Plant Growth by Reducing Biotic Stress" *Applied Sciences* 11, no. 5: 2282. https://doi.org/10.3390/app11052282

[4] Shorten, C., Khoshgoftaar, T.M. A survey on Image Data Augmentation for Deep Learning. *J Big Data* 6, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0

[5] H. Luo et al., "Normalized Avatar Synthesis Using StyleGAN and Perceptual Refinement," pp. 11657– 11667, 2021, doi: 10.1109/cvpr46437.2021.01149.

[6] Zhang, Yan, Shiyun Wa, Yutong Liu, Xiaoya Zhou, Pengshuo Sun, and Qin Ma. 2021. "High-Accuracy Detection of Maize Leaf Diseases CNN Based on Multi-Pathway Activation Function Module" Remote *Sensing* 13, no. 21: 4218. https://doi.org/10.3390/rs13214218

[7] Panigrahi, K.P., Das, H., Sahoo, A.K. and Moharana, S.C., 2020. Maize leaf disease detection and classification using machine learning algorithms. In *Progress in Computing, Analytics and Networking* (pp. 659-669). Springer, Singapore.

[8] Sibiya, Malusi, and Mbuyu Sumbwanyambe. 2019. "A Computational Procedure for the Recognition and Classification of Maize Leaf Diseases Out of Healthy Leaves Using Convolutional Neural Networks" *AgriEngineering* 1, no.1:119-131. https://doi.org/10.3390/agriengineering1010009

[9] B. K. Shah, A. Kumar, and A. Kumar, "Dog breed classifier for facial recognition using convolutional neural networks," Proc. 3rd Int. Conf. Intell. Sustain. Syst. ICISS 2020, pp. 508–513, 2020, doi: 10.1109/ICISS49785.2020.9315871.

[10] Hooker, Arthur L. "Corn and sorghum rusts." In *Diseases, Distribution, Epidemiology, and Control*, pp. 207-236. Academic Press, 1985.

[11] J, ARUN PANDIAN; GOPAL, GEETHARAMANI (2019), "Data for: Identification of Plant Leaf Diseases Using a 9-layer Deep Convolutional Neural Network", Mendeley Data, V1, doi: 10.17632/tywbtsjrjv.1

[12] Hussain, M., Bird, J.J., Faria, D.R. (2019). A Study on CNN Transfer Learning for Image Classification. In: Lotfi, A., Bouchachia, H., Gegov, A., Langensiepen, C., McGinnity, M. (eds) Advances in Computational Intelligence Systems. UKCI 2018. Advances in Intelligent Systems and Computing, vol 840. Springer, Cham. https://doi.org/10.1007/978-3-319

[13] Quang Hung Nguyen, Hai-Bang Ly, Lanh Si Ho, Nadhir Al-Ansari, Hiep Van Le, Van Quan Tran, Indra Prakash, Binh Thai Pham, "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil", *Mathematical Problems in Engineering*, vol. 2021, Article ID 4832864, 15 pages, 2021. https://doi.org/10.1155/2021/4832864

[14] Raschka, Sebastian, "Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning", https://doi.org/10.48550/arxiv.1811.12808

[15] Zhu, Qiuyu, Zikuang He, Tao Zhang, and Wennan Cui. 2020. "Improving Classification Performance of Softmax Loss Function Based on Scalable Batch-Normalization" *Applied Sciences* 10, no. 8: 2950. https://doi.org/10.3390/app10082950

[16] Xiaoling Xia, Cui Xu and Bing Nan, "Inception-v3 for flower classification," 2017 2nd International Conference on Image, Vision and Computing (ICIVC), 2017, pp. 783-787, doi: 10.1109/ICIVC.2017.7984661.

[17] Ruder, Sebastian, "An overview of gradient descent optimization algorithms", https://doi.org/10.48550/arxiv.1609.04747

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:06:50 UTC from IEEE Xplore. Restrictions apply.

Bulletin of Environmental Contamination and Toxicology
https://doi.org/10.1007/s00128-022-03572-w

Diurnal and Seasonal Variation of Atmospheric Particulate Matter and Trace Gases in Industrial Area of Delhi: A Study

Papiya Mandal[1] · Raju Sarkar[2] · Neel Kamal[3] · Manob Das[4] · Anubha Mandal[2]

Received: 23 August 2021 / Accepted: 24 June 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

The diurnal and seasonal variation of PM_{10}, SO_2, NO_2, NH_3 and water-soluble compounds were studied in Naraina industrial area; Delhi from January to December, 2017. It was observed that annual average concentrations of $PM_{10,}$ SO_2, NO_2, $NH_{3,}$ SO_4^{2-}, NO_3^-, NH_4^+,Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^- in the day time were 227 ± 91, 9 ± 5, 59 ± 22, 65 ± 15, 17.45 ± 5.14, 17.60 ± 4.94, 8.66 ± 2.94, 4.05 ± 1.08, 3.46 ± 0.91, 10.38 ± 4.48, 3.15 ± 0.99, 43.06 ± 5.20 and 0.50 ± 0.12 µg m^{-3}, respectively and night time were 320 ± 127, 14 ± 7, 82 ± 25, 83 ± 20, 22.64 ± 5.22, 21.66 ± 5.0, 11.81 ± 3.47, 3.29 ± 0.87, 3.02 ± 1.19, 7.55 ± 3.16, 2.49 ± 0.95, 31.86 ± 4.70 and 0.37 ± 0.12 µg m^{-3}, respectively. PM_{10} and sometimes NO_2 concentrations exceeded the Indian National Ambient Air Quality Standards. SO_2, and NH_3 concentrations were within the standard. The selected parameters varied from season to season. In the night time, selected parameters concentrations were high in comparison to day time might be due to formation of inorganic secondary particulate matters and low wind speed in the ambient air.

Keywords Trace gases · Particulate matter · Trace metals · Atmospheric pollutants · Delhi

In the present-day scenario, the characteristics of trace gases (SO_2, NO_2 and NH_3) and its changing characteristics to form atmospheric aerosols of sulphate (SO_4^{2-}), nitrates (NO_3^-) and ammonium (NH_4^+) are gaining notable interest to researchers especially in Indo Gangetic Plain (IGP) due to rapid urbanization, commercialization industrialization and growth of vehicular emissions this effect has been widely studied by Sen (2017), Das (2021), Shukla (2020) and Saud (2011). PM_{10} (particulate matter 10 µm or less in aerodynamic diameter) is the heterogeneous mixture of nano particles and liquid droplets. PM_{10} concentration is always alarmingly high and mostly exceeded four to six times of the standard of Indian National Ambient Air Quality Standards (NAAQS), of Central Pollution Control Board (CPCB-2009) in the urbanized Delhi city. PM_{10} has its own physical, chemical and thermodynamic properties and originates from both natural and anthropogenic sources reported by Perrino (2011).The exposure of PM_{10} is accountable with numerous persistent diseases and sometimes cancers to humans and generally classified as a human carcinogen this effect has been widely studied by Li (2018), Weber (2016) and Liu (2017). The various size of PM are accountable to climate change by interrupting the stability of earth's energy, hydrological circles, atmospheric circulation and abundance of green house and reactive trace gases reported by Poschl (2006). Water-soluble inorganic compounds and elements are the paramount components of PM_{10} and the elements ascendancy depends on numerous factors like meteorological parameters, geographical conditions and natural and anthropogenic sources of emission reported by Kliengchuay (2018) and (Kumari 2013). The common elements like sodium (Na), potassium (K), calcium (Ca), magnesium (Mg), manganese (Mn), aluminium (Al), and iron (Fe) are mostly present in the natural crust and directly emitted into the atmosphere from natural sources. The secondary aerosols are devised due to series of atmospheric reactions by two major particles formation processes reported by Wang (2016) and Leonardi (2020). The first operation is the nucleation process of low vapour pressure gases. The next operation is the condensation of low vapour pressure gases on the

✉ Papiya Mandal
papiya.mandal1942@gmail.com; p_mandal@neeri.res.in

[1] CSIR-NEERI, Zonal Centre, New Delhi 110 028, India

[2] Department of Civil Engineering, Delhi Technological University, Delhi 110 042, India

[3] CSIR- NEERI, Nehru Marg, Nagpur 440020, India

[4] Department of Geography, University of Gour Banga, Malda 732103, West Bengal, India

Published online: 25 July 2022

surface of the existing particles. Fine particles include both the accumulation and a major part of PM_{10}. The atmospheric nucleation processes may involve organic compounds, NH_3, water vapour, homogenous hetero molecular nucleation of HNO_3 and H_2SO_4 vapor formed due to oxidation of nitrogen oxides ($NOx = NO + NO_2$), SO_2 and OH radicals, present in the atmosphere. The nucleation process escalates both particle numbers and particle mass, while condensation guides only to an increase the particle mass and surface area. The several research are being accompanied to understand aerosol/particulate matter formation, its physical and chemical properties and their unfavorable effect to climate change and human health this effect has been widely studied by Shen et al. (2017), He (2018), Alemayehu (2020), Deng (2019), and Yang (2019). PM_{10} is the combination of water-soluble species, sea salt, mineral dust, organic carbon, elemental carbon, heavy metals etc. informed by Menon (2002). Aerosol has the capability to affect atmospheric radiation and disturb the balancing of climate forcing; nuclei cloud formation and photo-chemical reactions reported by Charlson (1992). The light extinction effects are influencing the global climate change and visibility of the people observed by Tsai (2003). Trace gases is the mixture of conglomerate gases and by volume; concentration is less or equal to one percent only. Initially, homogeneous trace gases reached upto supersaturated state through adiabatic expansion or mixing with cool air and then converted to particulate matter (PM) by the oxidation process reported by Friedlander (2000). The emission of trace gases is metamorphosed to secondary inorganic PM due to photochemical reactions. The primary precursors are metamorphosed to sulphate (SO_4^{2-}), nitrate (NO_3^-) and ammonium (NH_4^+) from trace gases. This heterogeneous complex process influences to accumulate PM in the ranges of 0.1–2.5 µm accepted by Perrino (2011). The secondary inorganic aerosols (SO_4^{2-}, NO_3^- and NH_4^+) are the predominant components of $PM_{2.5}$ (fine inhalable tiny particles that are concern to people's health when exceeds the permissible level). A $PM_{2.5}$ concentration is accounted to 45.9% of total mass approximately, reported by Li (2015). The rate of conversion of trace gases in to inorganic secondary particles depends on factors like precursor concentrations; reactive gaseous concentrations (eg., ozone, hydroxyl radical, nitrate radical, peroxy radicals, hydrogen peroxide and organic carbon etc.) and meteorological parameters (eg., temperature, humidity, wind speed, mixing height, solar insolation etc. observed by Gupta (2003) and Kumari (2013). A few studies were reported the diurnal and seasonal variation of PM_{10} along with trace gases (SO_2, NO_2, and NH_3) and water-soluble inorganic compound in industrial area of Delhi, India. The main objective of the present study is to knowhow the diurnal and seasonal variation of PM_{10}, SO_2, NO_2, and NH_3 and the contribution of water-soluble species in PM_{10} in Delhi.

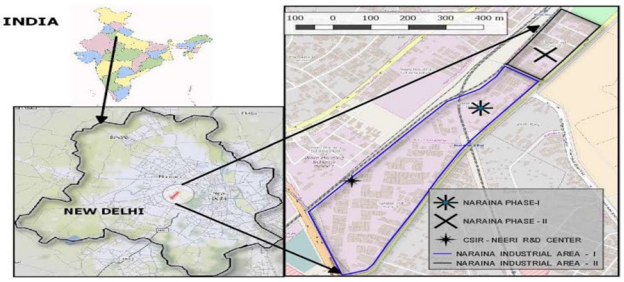


Fig. 1 Sampling Location map

Materials and Methods

The sampling was accomplished at Council of Scientific and Industrial Research (CSIR), Development Centre, Naraina industrial area (28°12′–28°63′N, 75°50′–77°23′E) in Delhi is shown in Fig. 1. The selected atmospheric parameters PM_{10}, SO_2, NO_2, and NH_3 samples were collected for 1 year (January to December 2017) for 6–8 days in a month on 12 hourly bases (6 am to 6 pm and 6 pm to 6 am) round the clock. Total 100 nos of PM_{10} samples (100 nos days and 100 nos night samples) were collected by Fine Dust Sampler (APM 550, make Envirotech Pvt. Ltd. India) at the flow rate of 16.7 litre per minute. APM 550 without cascade impactor may be use for collection of PM_{10} samples and PM_{10} was collected in the same procedure in the year 2017. Initially, quartz filter papers were prebaked (at 550°C, for 6 h) in the muffle furnace to detach the organic impurities if any. Quartz filter papers have high resistance to chemicals, and do not absorb dioxides of NO_x and SO_x. It remains uninfluenced by humidity and is easily sterilized. This could be the reason for uses of quartz filter paper in air pollution analysis purposes. The filter papers were kept in a desiccator and maintained humidity 45% ± 5% and temperature 25 ± 2°C for 24 h to remove moisture before weighing. The initial and final weight of the filter papers were measured by using Greifensee, Switzerland Mettler–Toledo balance (resolution: ± 1 µg). PM_{10} concentrations (in µg m^{-3}) were calculated from the difference between the initial and final weight of filter paper divided by the total volume of air passed. For QA/QC of the analysis, one tenth of the samples were weighed again by adopting the same procedure and calculated PM_{10} concentration. The concentration of PM_{10} of the reweighed filter paper was within the 5% range of the calculated value of PM_{10} concentration (which is acceped as correct value). The filter papers were stored into the deep freezer at (− 20°C) before further analysis of water-soluble and trace metals compounds. A set of three impinges filled with 20 ml of selected absorbing media in series were retained into the gaseous sampler (APM 411TE, make Envirotech Pvt. Ltd. India) to collect ambient SO_2, NO_2, and NH_3 samples at

a flow rate of 1.0, 0.5 and 0.5 lpm, respectively. Initially, the absorbing media eg., potassium-tetra-chloromercurate (for SO_2), sodium-hydroxide and sodium- arsenate (for NO_2) and diluted sulphuric acid (for NH_3) were used for trapping ambient gaseous pollutants. The selected gaseous parameters (SO_2, NO_2 and NH_3) were analysed by wet chemical methods like Improved West and Gaeke, Modified Jacob–Hochheiser and Indophenol Blue Method respectively to measure ambient SO_2, NO_2, and NH_3 (in µg m^{-3}). Similar procedure, three blank impingers were also filled with 20 ml of absorbing media and analyzed with the collected samples to maintained QA and QC. The above mentioned methods are reference methods of CPCB. Standard calibration graphs are also prepared for all the selected gaseous parameters and calibration factors were used to find out SO_2, NO_2, and NH_3 concentrations (in µg m^{-3}). The minimum detection limit of concentrations for SO_2 is 2 µg m^{-3}, for NO_2 is 9 µg m^{-3} and for NH_3 is 2 µg m^{-3} respectively.

The weekly based collected diurnal (day and night) filter papers of PM_{10} (total 50 nos.) were extracted for ninety minutes in ultrasonic extractor with 40 ml deionized (ultrapure) water for analysis of common cations eg., sodium (Na^+), potassium (K^+), ammonium (NH_4^+), magnesium (Mg^{2+}) calcium (Ca^{2+}) and anions eg., sulphate (SO_4^{2-}), nitrate (NO_3^-), chloride (Cl^-) and fluoride (F^-). After ultrasonic extraction, samples were filtered through 0.22 micron filter paper to separate insoluble impurities if any, then transferred to plastic bottles and kept in the refrigerator. The cations and anions were investigated in a sophisticated instrument like Ion Chromatograph (DIONEX-ICS-5000, Make: ThermoFisher Scientific, USA). The concentrations of SO_4^{2-}, NO_3^-, Cl^- and F^- were determined by using an Ion Pac-AS11-HC analytical column (4 × 250 mm, Dionex, USA) with a guard column (IonPac AG11-HC, 4 × 50 mm, Dionex, USA). Na^+, K^+, NH_4^+, Mg^{2+} and Ca^{2+} were determined using a separation column (IonPac CS17-HC, 4 × 250 mm, Dionex, USA) with a guard column (IonPac CG17-HC, 4 × 50 mm, Dionex, USA), suppressor CSRS-300 (4 mm, Dionex, USA) and 5 mM of MSAas the eluent.

The annual average rainfall in Delhi (the capital of India) in the year 2017 was 617 mm and in general rainfall varies from 600 to 800 mm, occurs mostly during the monsoon season in the months of July and August of every year. The meteorological parameters eg., wind speed and wind direction were collected for a year 2017 from CPCB website (https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/data) to understand the climatic pattern in Delhi in the year 2017.The diurnal basis variation of meteorological parameters in Delhi in the year 2017 is presented in Fig. 2. Wind direction was predominant in the west, northwest, and west-northwest directions especially in winter, premonsoon and postmonsoon seasons and east and east-southeast directions in the monsoon season. It was observed that in the year 2017 wind speed is more in the day time as compared to night time.

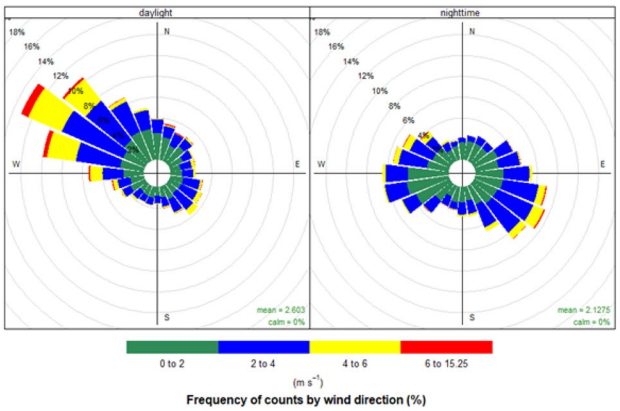


Fig. 2 Dinual variation of Meteorological conditions of Delhi in the year 2017

Results and Discussion

The average diurnal and monthly variation of PM_{10} and trace gases (SO_2, NO_2 and NH_3) at Naraina, Delhi in the year 2017 is presented in Fig. 3a–d. It was observed that selected parameters were varied from season to season, month to month, and even in day and night basis too. The concentrations of trace gases in the atmosphere are mainly due to industrial emission, vehicular pollution and biomass burning especially during the postmonsoon and winter season. The decrease of wind speed in the night time reduces the dispersion of the pollutants from one to another location might be the reason to increase the concentrations of atmospheric pollutants. Trace gases are actively played role to increase PM_{10} concentration in the ambient air. PM_{10} concentration exceeded (NAAQS, 2009) 4 to 6 times (permissible limit of annual mean 60 µg m^{-3}). The higher concentrations of PM_{10}, SO_2, NO_2 and NH_3 in day and night time basis was observed in the month of December 2017 were found to be 366 ± 48, 17 ± 2, 104 ± 15, 77 ± 12, and 480 ± 47, 24 ± 4, 120 ± 13and 102 ± 17 µg m^{-3}, respectively. The lowest concentrations of PM_{10}, SO_2, NO_2 and NH_3 in day and night time basis were observed in July 2017 were 92 ± 18, 3 ± 1, 28 ± 2, 55 ± 10 and 124 ± 29, 4 ± 1, 41 ± 8, and 67 ± 12 µg m^{-3}, respectively. The average concentrations of PM_{10}, SO_2, NO_2 and NH_3 in day and night basis during winter (265 ± 72, 13 ± 3, 81 ± 14 and 62 ± 9 and 392 ± 80, 21 ± 3, 107 ± 9, and 77 ± 15 µg m^{-3}), premonsoon (213 ± 64, 5 ± 1, 53 ± 13 and 60 ± 10 and 286 ± 104, 9 ± 3, 72 ± 16, and 74 ± 7 µg m^{-3}), Monsoon (132 ± 52, 4 ± 3, 32 ± 4 and 67 ± 13 and 182 ± 88, 7 ± 5, 51 ± 10, and 82 ± 16 µg m^{-3}) and postmonsoon (318 ± 27, 12 ± 2, 66 ± 7, 77 ± 14, and 435 ± 25, 20 ± 5, 93 ± 6, and

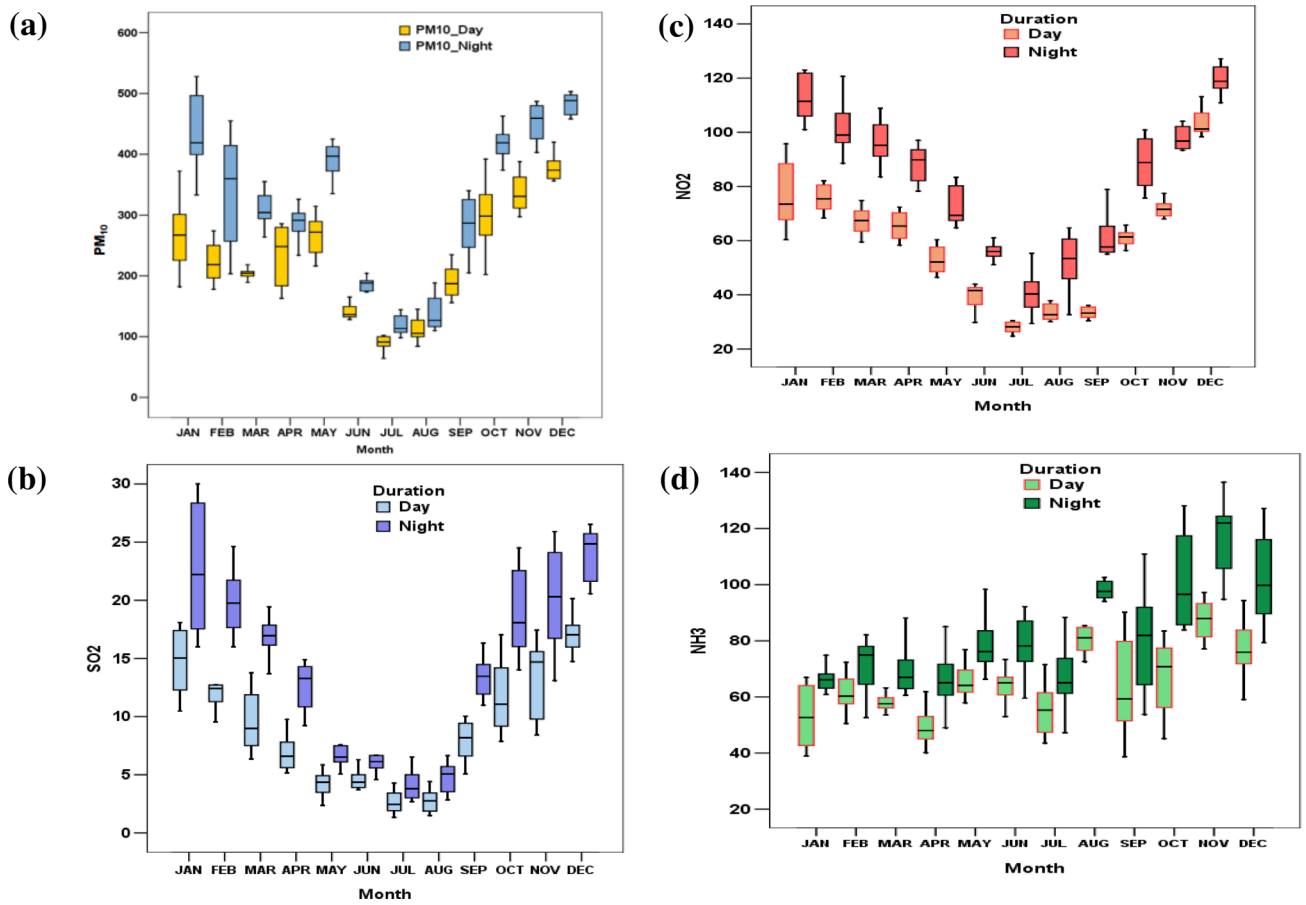


Fig. 3 Diurnal and monthly variation of **a** PM_{10}, **b** SO_2, **c** NO_2 and **d** NH_3 in Delhi (2017)

109 ± 11 μg m^{-3}), respectively. The concentrations of SO_2, and NH_3 were observed within the permissible limit of NAAQS, but NO_2 concentration sometimes exceeded the permissible limit in NAAQS espcially during the postmonsoon and winter season. The diurnal and seasonal variation of water-soluble compounds (SO_4^{2-}, NO_3^-, NH_4^+, Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^-) is presented in Table 1. Trace metals, cation and anion were measured in four seasons like winter (January–March and December), premonsoon (April–June), monsoon (July−September) and Postmonsoon (October–November) in the year 2017. The highest concentrations of selected parameters were observed in postmonsoon season followed by winter, premonsoon, and monsoon season. In the postmonsoon, the concentrations of SO_4^{2-}, NO_3^-, NH_4^+, Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^- in day and night time were found to be 22.83 ± 5.57, 23.73 ± 4.27, 8.44 ± 2.70, 3.84 ± 0.55, 3.12 ± 0.76, 8.0 ± 1.65, 3.49 ± 0.73, 42.39 ± 4.41, 0.48 ± 0.20 and 27.18 ± 5.30, 26.63 ± 3.47, 13.52 ± 4.35, 3.34 ± 0.73, 3.63 ± 1.75, 6.32 ± 1.20, 2.95 ± 0.92, 28.72 ± 6.52, 0.35 ± 0.15 μg m^{-3}, respectively. In the winter season, 19.88 ± 3.59, 20.22 ± 2.40, 10.29 ± 1.52, 4.62 ± 1.27, 3.69 ± 0.93, 15.19 ± 3.55, 3.67 ± 0.84, 45.97 ± 6.32, 0.56 ± 0.09 and 26.33 ± 2.84, 25.11 ± 3.25, 13.98 ± 2.32, 3.55 ± 0.77, 3.07 ± 1.25, 10.77 ± 3.0, 2.71 ± 0.94, 34.97 ± 3.97, 0.46 ± 0.10 μg m^{-3}, respectively. In the premonsoon season, in day and night time were found to be 16.35 ± 2.63, 11.98 ± 1.96, 7.56 ± 1.45, 3.63 ± 0.88, 4.11 ± 0.67, 9.29 ± 2.40, 3.35 ± 0.82, 42.02 ± 3.39, 0.48 ± 0.09 and 21.06 ± 1.98, 16.96 ± 3.27, 9.58 ± 1.12, 3.41 ± 0.99, 3.42 ± 0.67, 5.88 ± 1.86, 2.73 ± 0.84, 33.12 ± 2.54, 0.29 ± 0.06 μg m^{-3}, respectively. In the monsoon season in day and night time were found to be 11.73 ± 1.35, 15.63 ± 2.19, 7.73 ± 4.56, 3.85 ± 1.05, 2.71 ± 0.54, 6.66 ± 2.61, 2.04 ± 0.56, 40.65 ± 4.14, 0.46 ± 0.09 and 16.29 ± 1.55, 18.44 ± 1.80, 10.03 ± 3.61, 2.79 ± 0.82, 2.16 ± 0.47, 5.74 ± 1.80, 1.66 ± 0.61, 28.54 ± 2.04, 0.34 ± 0.07 μg m^{-3}, respectively. The highest and lowest concentration of trace metals in the atmosphere were observed in the postmonsoon and monsoon season, respectively. In general the formation of SO_4^{2-}, NO_3^-, and NH_4^+ in the atmosphere depends on atmospheric oxidants/photochemistry reaction. It was also noticed that oxidants/photochemistry reaction occurred maximum in the postmonsoon season followed by winter, premonsoon and monsoon season. In the

Table 1 Diurnal and seasonal variation of atmospheric water soluble compounds (in μg m^{-3})

Parameter	Winter (Jan- Mar, 2017, & Dec 2017)		Premonsoon (Apr- Jun, 2017)		Monsoon (Jul- Sep, 2017)		Postmonsoon (Oct- Nov, 2017)		Annual Average Day	Annual Average Night
	Day	Night	Day	Night	Day	Night	Day	Night		
SO_4^{2-}	19.88 ± 3.59	26.33 ± 2.84	16.35 ± 2.63	21.06 ± 1.98	11.73 ± 1.35	16.29 ± 1.55	22.83 ± 5.57	27.18 ± 5.30	17.45 ± 5.14	22.64 ± 5.22
NO_3^-	20.22 ± 2.40	25.11 ± 3.25	11.98 ± 1.96	16.96 ± 3.27	15.63 ± 2.19	18.44 ± 1.80	23.73 ± 4.27	26.63 ± 3.47	17.60 ± 4.94	21.66 ± 5.0
NH_4^+	10.29 ± 1.52	13.98 ± 2.32	7.56 ± 1.45	9.58 ± 1.12	7.73 ± 4.56	10.03 ± 3.61	8.44 ± 2.70	13.52 ± 4.35	8.66 ± 2.94	11.81 ± 3.47
Na^+	4.62 ± 1.27	3.55 ± 0.77	3.63 ± 0.88	3.41 ± 0.99	3.85 ± 1.05	2.79 ± 0.82	3.84 ± 0.55	3.34 ± 0.73	4.05 ± 1.08	3.29 ± 0.87
K^+	3.69 ± 0.93	3.07 ± 1.25	4.11 ± 0.67	3.42 ± 0.67	2.71 ± 0.54	2.16 ± 0.47	3.12 ± 0.76	3.63 ± 1.75	3.46 ± 0.91	3.02 ± 1.19
Ca^{2+}	15.19 ± 3.55	10.77 ± 3.0	9.29 ± 2.40	5.88 ± 1.86	6.66 ± 2.61	5.74 ± 1.80	8.00 ± 1.65	6.32 ± 1.20	10.38 ± 4.48	7.55 ± 3.16
Mg^{2+}	3.67 ± 0.84	2.71 ± 0.94	3.35 ± 0.82	2.73 ± 0.84	2.04 ± 0.56	1.66 ± 0.61	3.49 ± 0.73	2.95 ± 0.92	3.15 ± 0.99	2.49 ± 0.95
Cl^-	45.97 ± 6.32	34.97 ± 3.97	42.02 ± 3.39	33.12 ± 2.54	40.65 ± 4.14	28.54 ± 2.04	42.39 ± 4.41	28.72 ± 6.52	43.06 ± 5.20	31.86 ± 4.70
F^-	0.56 ± 0.09	0.46 ± 0.10	0.48 ± 0.09	0.29 ± 0.06	0.46 ± 0.09	0.34 ± 0.07	0.48 ± 0.20	0.35 ± 0.15	0.50 ± 0.12	0.37 ± 0.12

postmonsoon and winter season this could be possible due to the favorable meteorological condition such as atmospheric stability, low to average mixing height, slow dispersion etc. as compared to premonsoon and monsoon seasons observed by Mandal (2014). The anion and cation like SO_4^{2-}, NO_3^-, NH_4^+ had significant variation whereas trace metals like Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^- had no notable variation. The present study observation is comparable with another study carried out in Delhi by Sharma (2014). Sometimes it was observed that atmospheric NO_3^- concentration was higher as compared to SO_4^{2-} concentration might be due to local emission of NO_x and finally converted to NO_3^-. According to Li (2006), the concentration of NO_3^- remains relatively low due to higher volatility of particulate NO_3^- during premonsoon season as compared to other seasons. A higher concentration of relative humidity might be due to dissolutions of NH_3 that increases NH_4^+ functions reported by Trebs (2004) and Ianniello (2010). Cl^- has light-absorbing properties and plays an important role in the boundary layer chemistry and climate change informed by Wang (2016). Cl^- concentration was abnormally high throughout the year might be due to re-suspension of windblown dust, burning of solid waste informed by Li et al. (2012) and aerosol mass transportation from marine location. In day and night time, Cl^- concentration varied from 13% to 39% with an average 22%, and 5 to 26 with an average of 12%. F^- had no significant variation, might be due to non-availably of source. The presence of concentrations of atmospheric Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^- were might be due to re-suspension of windblown dust epecially in the premonsoon and monsoon season. Potassium (K^+) concentration was observed high in the postmonsoon and winter season in Delhi NCR due to biomass/ stubble burning. It was reported by Penner (1995) that K^+ is the marker of biomass burning. In the present study, trace metals contribution in PM_{10} in day and night varied from 18% to 59% with an average percentage of 32% and 9% to 35% with an average percentage of 18%, respectively. The correlation coefficient and linear regression equation of SO_4^{2-} and NO_3^- with cations (Na^+, K^+, Ca^{2+}, Mg^{2+}, and NH_4^+) is shown in Table 2. SO_4^{2-} had a strong correlation coefficient (r = 0.85) with both K^+ and NH_4^+, indicating thepossitibility to form inorganic compounds like K_2SO_4 and $(NH_4)_2SO_4$ and weak correlation with Na^+, Ca^{2+} and Mg^{2+}, indicating a comparatively less formation of Na_2SO_4, $CaSO_4$, and $MgSO_4$ in the atmosphere. NO_3^- had a strong correlation coefficient (r = 0.81) with NH_4^+ and K^+ (r = 0.70) suggested the atmospheric formation of NH_4NO_3 and KNO_3. The moderate correlation with Na^+ and Mg^{2+} indicated that less formation of $NaNO_3$ and $Mg(NO_3)_2$. The poor correlation with Ca^{2+} indicated the low formation of $Ca(NO_3)_2$. NH_4^+ is playing an active role to neutralize the acidic anion like SO_4^{2-} and NO_3^- in the atmosphere. The acidic species SO_4^{2-} and NO_3^- were also neutralized by the natural trace metals like

Table 2 The correlation and regression of water soluble compounds

Anion	Correlation and regression with cation				
	Na^+	K^+	Ca^{+2}	Mg^{+2}	NH_4^+
SO_4^{-2}	r (SO_4^{-2}: Na^+) = 0.45 Reg SO_4^{-2} = 0.06Na^+ + 2.45	r (SO_4^{-2}: K^+) = 0.85 Reg SO_4^{-2} = 0.18 K^+-0.18	r (SO_4^{-2}: Ca^{+2}) = 0.38 Reg SO_4^{-2} = 0.18Ca^{+2} + 4.25	r (SO_4^{-2}: Mg^{+2}) = 0.42 Reg SO_4^- = 0.08 Mg^{+2} + 1.55	r (SO_4^{-2}: NH_4^+) = 0.85 Reg SO_4^{-2} = 0.59 NH_4^+-2.03
	R^2 = 0.21	R^2 = 0.72	R^2 = 0.14	R^2 = 0.18	R^2 = 0.73
NO_3^-	r (NO_3^-: Na^+) = 0.59 Reg NO_3^- = 0.12Na^+ + 1.78	r (NO_3^-: K^+) = 0.70 Reg NO_3^- = 0.22 K^+-0.16	r (NO_3^-: Ca^{+2}) = 0.42 Reg NO_3^- = 0.28Ca^{+2} + 3.2	r (NO_3^-: Mg^{+2}) = 0.48 Reg NO_3^- = 0.13 Mg^{+2} + 1.0	r (NO_3^-: NH_4^+) = 0.81 Reg NO_3 = 0.80NH_4^+-3.57
	R^2 = 0.35	R^2 = 0.50	R^2 = 0.17	R^2 = 0.23	R^2 = 0.66

r Correlation Coefficient, *Reg* Regression, R^2 Regression Coeffient

Na^+, K^+, Ca^{2+}, and Mg^{2+}. Generally, NH_4^+ is converted to NH_3 at high temperature (> 25°C) and low RH reported by Sharma (2007). Night time, PM_{10} and inorganic compounds concentrations were higher as compared to day time, might be due to the formation of secondary inorganic particulate matter in the atmosphere. In the night time, the increased concentrations of PM_{10}, SO_2, NO_2, NH_3, SO_4^{2-}, NO_3^- and NH_4^+ were 41%, 56% 39%, 28%, 30%, 23% and 36%, respectively. Moreover, in the night time, the decreased concentrations of Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^- were 19%, 13%, 27%, 21%, 26% and 26%, respectively. It was reported, that annual average increased night time concentration of $PM_{2.5}$, (part of PM_{10}) was 23% higher than those observed for daytime in urban sites of central India observed by Pervez S (2020). Our observed data is comparable with previously reported data of many researchers also. Lower night time temperatures; result in increased domestic heating, and night time inversions accumulate air pollutants Pervez S (2020). The variation of industrial pollution, vehicular pollution, meteorological parameters and atmospheric mixing ratios are the governing parameters to control the diurnal variation in the ambient air Sharma (2014) and Mandal (2014). The contaminated particulater matter due to sulphate, nitrate and ammonium salt has adverse health effect like reduction of lung function, aggravation of asthmatic symptoms, and increasing risk of emergency department visits, hospitalizations, and death in people in the extreme situation.

It can be concluded that trace gases (SO_2, NO_2, and NH_3) played an important role to form inorganic secondary particulate matter and increase the atmospheric PM_{10} concentrations. Day time, PM_{10}, SO_2, NO_2, and NH_3 concentrations were lower as compared to night time, might be due to the conversion of trace gases to inorganic particulate matter in the night time and minimum wind speed in the night time as compared to day time. The concentrations of secondary forms (SO_4^{2-}, NO_3^- and NH_4^+) were higher in postmonsoon and winter season might be due to the effect of biomass burning in Delhi and adverse meteorological conditions. SO_4^{2-} and NO_3^- were highly correlated with NH_4^+ and K^+ as well. Na^+, K^+, Ca^{2+}, Mg^{2+}, Cl^- and F^- varied from season to season. Cl^- was only divergent compound whose concentration was abnormally high throughout the year might be due to the burning of solid waste/garbage through out the year and transportation of aerosol from marine and other location to Delhi. The emissions of SO_2, NO_2, and NH_3 and their secondary pollutants have the major role to increase the concentration of contamined particulate matter and may cause public health effects.

Acknowledgements The authors are grateful to the Council of Scientific and Industrial Research (CSIR) for sponsoring the research under CSIR-EMPOWER scheme.

References

Alemayehu YA, Asfaw SL, Terfie TA (2020) Exposure to urban particulate matter and its association with human health risks. Environ Sci Pollut Res Int 27(22):27491–27506. https://doi.org/10.1007/s11356-020-09132-1

Charlson RJ, Schwartz SE, Hales JM, Cess RD, Coakley JA, Hansen JE, Hofmann DJ (1992) Climate forcing by anthropogenic aerosol. Science 255:423–430. https://doi.org/10.1126/science.255.5043.423

Das M, Das A, Ghosh S, Sarkar R, Saha S (2021) Spatio-temporal concentration of atmospheric particulate matter (PM2.5) during pandemic: a study on most polluted cities of indo-gangetic plain. Urban Climate 35:100758. https://doi.org/10.1016/j.uclim.2020.100758

Deng Q, Deng L, Miao Y, Guo X, Li Y (2019) Particle deposition in the human lung: health implications of particulate matter from different sources. Environ Res 169:237–245. https://doi.org/10.1016/j.envres.2018.11.014

Friedlander SK (Second Edition, 2000) Smoke, Dust, and Haze: Fundamentals of aerosol dynamics. Chapter-10-gas-to-particle-conversion. https://www.globalspec.com/reference/77570/203279

Gupta A, Kumar R, Kumari KM, Srivastava SS (2003) Measurement of NO_2, HNO_3, NH_3 and SO_2 and related particulate matter at a rural site in Rampur, India. Atmos Environ 37:4837–4846. https://doi.org/10.1016/j.atmosenv.2003.07.008

He H, Liang XZ, Wuebbles DJ (2018) Effects of emissions change, climate change and long-range transport on regional modeling of future US particulate matter pollution and speciation. Atmos Environ 179:166–176. https://doi.org/10.1016/j.atmosenv.2018.02.020

Ianniello A, Spataro F, Esposito G, Allegrini I, Rantica E, Ancora MP et al (2010) Occurrence of gas phase ammonia in the area of Beijing (China). Atmos Chem Phys 10:9487–9503. https://doi.org/10.5194/acp-10-9487-2010

Kliengchuay W, Cooper Meeyai A, Worakhunpiset S, Tantrakarnapa K (2018) Relationships between meteorological parameters and particulate matter in Mae Hong Son Province, Thailand. Int J Environ Res Public Health 15(12):2801. https://doi.org/10.3390/ijerph15122801

Kumari KM, Satsangi A, Pachauri T, Singla V, Lakhani A (2013) Water soluble ionic species in atmospheric aerosols: concentrations and sources at Agra in the Indo-Gangetic Plain (IGP). Aerosol and Air Qual Res 13:1877–1889. https://doi.org/10.4209/aaqr.2012.08.0227

Leonardi A, Ricker HM, Gale AG, Ball BT, Odbadrakh TT, Shields GC, Navea JG (2020) Particle formation and surface processes on atmospheric aerosols: a review of applied quantum chemical calculations. Int J Quantum Chem 2020:e26350. https://doi.org/10.1002/qua.26350

Li Y, Schwab JJ, Demerjian KL (2006) Measurements of ambient ammonia using a tunable diode laser absorption spectrometer: characteristics of ambient ammonia emissions in an urban area of New York City. J Geophys Res Atmos. https://doi.org/10.1029/2005JD006275

Li G, Lei W, Bei N, Molina LT (2012) Contribution of garbage burning to chloride and $PM_{2.5}$ in Mexico City. Atmos Chem Phys 12:8751–8761. https://doi.org/10.5194/acp-12-8751-2012

Li X, Wang H, Tian M, Chang Q, Cao J, Yang F, Ma Y, He K (2015) Chemical composition and light extinction contribution of PM2.5 in urban Beijing for a 1-year period. Aerosol Air Qual Res 15(6):2200–2211. https://doi.org/10.4209/aaqr.2015.04.0257

Li T, Hu R, Chen Z, Li Q, Huang S, Zhu Z, Zhou LF (2018) Fine particulate matter (PM2.5): the culprit for chronic lung diseases in China. Chronic Dis Transl Med 4(3):176–186. https://doi.org/10.1016/j.cdtm.2018.07.002

Liu S, Zhou Y, Liu S, Chen X, Zou W, Zhao D et al (2017) Association between exposure to ambient particulate matter and chronic obstructive pulmonary disease: results from a cross-sectional study in China. Thorax 72(9):788–795. https://doi.org/10.1136/thoraxjnl-2016-208910

Mandal P, Saud T, Sarkar R, Mandal A, Sharma SK, Mandal TK, Bassin JK (2014) High seasonal variation of atmospheric C and particle concentrations in Delhi, India. Environ Chem Lett 12:225–230. https://doi.org/10.1007/s10311-013-0438-y

Menon S, Hansen JE, Nazarenko L, Luo Y (2002) Climate effects of black carbon aerosols in China and India. Science 297:2250–2253. https://doi.org/10.1126/science.1075159

Penner JE, Charlson RJ, Heintzenberg J (1995) Carbonaceous aerosol influencing atmospheric radiation: black carbon and organic carbon. Aerosol Forcing of Climate 91–108. https://www.osti.gov/servlets/purl/10118242

Perrino C, Tiwari S, Catrambone M, Torre SD, Rantica E, Canepari S (2011) Chemical characterization of atmospheric PM in Delhi, India, during different periods of the year including Diwali festival. Atmos Poll Res 2(4):418–427. https://doi.org/10.5094/APR.2011.048

Pervez S, Sahu RK, Tripathi M, Bano S, Matawle JL, Tiwari S, Deb MK, Pervez YF (2020) Assessment and evaluation of ambient PM2.5 in relation to its health effects in mineral-based coal-fired areas. Geofizika 37(1):67–90. https://doi.org/10.15233/gfz.2020.37.1

Poschl U (2006) Atmospheric aerosols: composition, transformation, climate and health effects. Angewandte Chemie Int Edition 44(46):7520–7540. https://doi.org/10.1002/anie.200501122

Saud T, Mandal TK, Gadi R, Singh DP, Sharma SK, Saxena M, Mukherjee A (2011) Emission estimates of particulate matter (PM) and trace gases (SO_2, NO and NO_2) from biomass fuels used in rural sector of Indo-Gangetic Plain. India Atmos Environ 45(32):5913–5923. https://doi.org/10.1016/j.atmosenv.2011.06.031

Sen A, Abdelmaksoud AS, Ahammed YN, Banerjee T, Bhat MA, Chatterjee A et al (2017) Variations in particulate matter over indo-gangetic plains and indo-himalayan range during four field campaigns in winter monsoon and summer monsoon: role of pollution pathways. Atmos Environ 154:200–224. https://doi.org/10.1016/j.atmosenv.2016.12.054

Sharma M, Kishore S, Tripathi SN, Behera SN (2007) Role of atmospheric ammonia in the formation of inorganic secondary particulate matter: A study at Kanpur, India. J Atmos Chem 58:1–17. https://doi.org/10.1007/s10874-007-9074-x

Sharma SK, Mandal TK, Saxena M, Rashmi S, Datta A, Saud T (2014) Variation of OC, EC, WSIC and trace metals of PM_{10} in Delhi. J Atmos Solar-Terrestrial Phys 113:10–22. https://doi.org/10.1016/j.jastp.2014.02.008

Shen L, Mickley LJ, Murray LT (2017) Influence of 2000–2050 climate change on particulate matter in the United States: results from a new statistical model. Atmos Chem Phys 17(6):4355–4367. https://doi.org/10.5194/acp-17-4355-2017

Shukla K, Kumar P, Mann GS, Khare M (2020) Mapping spatial distribution of particulate matter using kriging and inverse distance weighting at supersites of megacity Delhi. Sustain Cities Soc 54:101997. https://doi.org/10.1016/j.scs.2019.101997

Trebs I, Meixner FX, Slanina J, Otjes R, Jongejan P, Andreae MO (2004) Real-time measurements of ammonia, acidic trace gases and water-soluble inorganic aerosol species at a rural site in the Amazon Basin. Atmos Chem Phys 4(4):967–987. https://doi.org/10.5194/acp-4-967-2004

Tsai YI, Lin YH, Lee SZ (2003) Visibility variation with air qualities in the metropolitan area in southern Taiwan. Water Air Soil Pollut 144:19–40

Wang D, Zhou B, Fu Q, Zhao Q, Zhang Q, Chen J et al (2016) Intense secondary aerosol formation due to strong atmospheric photochemical reactions in summer: observations at a rural site in eastern Yangtze River Delta of China. Sci Total Environ 571:1454–1466. https://doi.org/10.1016/j.scitotenv.2016.06.212

Weber SA, Insaf TZ, Hall ES, Talbot TO, Huff AK (2016) Assessing the impact of fine particulate matter ($PM_{2.5}$) on respiratory-cardiovascular chronic diseases in the New York City metropolitan area using hierarchical bayesian model estimates. Environ Res 151:399–409. https://doi.org/10.1016/j.envres.2016.07.012

Yang Y, Ruan Z, Wang X, Yang Y, Mason TG, Lin H, Tian L (2019) Short-term and long-term exposures to fine particulate matter constituents and health: a systematic review and meta-analysis. Environ Pollut 247:874–882. https://doi.org/10.1016/j.envpol.2018.12.060

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/362716984

DNS-Tunnet: A Hybrid Approach for DNS Tunneling Detection

Conference Paper · August 2022

DOI: 10.1109/CTISC54888.2022

CITATIONS
0

READS
19

4 authors, including:

Anushka Lal
Delhi Technological University
4 PUBLICATIONS **13** CITATIONS
SEE PROFILE

Ankit Kumar
Delhi Technological University
3 PUBLICATIONS **2** CITATIONS
SEE PROFILE

All content following this page was uploaded by Anushka Lal on 16 August 2022.

The user has requested enhancement of the downloaded file.

DNS-Tunnet: A Hybrid Approach for DNS Tunneling Detection

Anushka Lal
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
anushkalal_2k18co083@dtu.ac.in

Ankit Prasad
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
ankitprasad_2k18co073@dtu.ac.in

Ankit Kumar
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
ankitkumar_2k18co071@dtu.ac.in

Shailender Kumar
Department of Computer Science and Engineering
Delhi Technological University
Delhi, India
shailenderkumar@dce.ac.in

***Abstract*—Domain Name Service (DNS) is a well-trusted protocol for translating domain name to IP address. However, cybercriminals have found new ways to launch attacks through this medium such as DNS Tunneling, by creating covert control and command channels to tunnel payloads. Most of the previous studies extract hand-crafted features from DNS queries and train them individually on state-of-the-art methods. In this study, we have not only explored the standard Deep Learning approach but have also used the integration of Convolutional Neural Networks (CNN) and Support Vector Machine (SVM) in our proposed DNS-Tunnet architecture. This technique utilizes both automatic feature extraction property of CNN and the precise classification property of SVM. For comparison purposes, we have compared performance of DNS-Tunnet with ensemble classifiers for which we prepared two different datasets through feature extraction. Our experiment concludes that DNS-Tunnet outperforms all other models both in terms of performance and efficiency.**

***Index Terms*—Convolutional Neural Network, Deep Learning, DNS Tunneling Detection, DNS-Tunnet, Domain Name System, Ensemble Learning, Support Vector Machines**

I. Introduction

Domain Name System (DNS) is a decentralized and hierarchical system that maps human-legible domain names (e.g., facebook.com) into machine-legible IP addresses (e.g., 157.240.3.35) [1]. Initially, the DNS protocol was not intended for data transfer, hence was not closely monitored by security systems. However lately, bad actors have started exploiting this vulnerability. For instance, recently uncovered government sponsored Project Sauron, used DNS tunneling technique to carry out data exfiltration [2]. A survey by Neustar International Security Council in September 2021, also claims that 72% of participating institutes faced DNS attacks in the past 12 months [3].

The fundamental aspect of exploiting DNS by bad actors is to have a registered domain name (e.g. maliciousssite.com) and its associated name server (NS) under its control, such that whenever the compromised host sends a DNS query of the form "www.sensitivedata.maliciousssite.com" (sensitivedata: the sensitive data encoded by the malware present inside the host) it reaches the bad actor through NS for "maliciousssite.com" where the sensitive data is finally extracted by the bad actor and a fake genuine response is returned to evade detection.

DNS Tunneling detection can be majorly divided into two categories: traffic analysis [4]-[6] and payload analysis [7]-[9]. Traffic analysis is the assessment of DNS traffic based on the count and frequency of requests. Payload analysis, on the other hand, is the assessment of the payload in a single DNS request, such as domain length, character ratio or count of letters, numbers and special characters.

Several works in the past for this domain have been in the application of supervised and unsupervised machine learning techniques that analyze the DNS query for any inserted payloads. While the performance of these models is acceptable, it does not eliminate the need to extract hand-crafted features from the query, which is overall a very time-consuming and complex process. The Deep Learning approach helps solve this limitation, and is the motivation behind this study.

Our Contribution:

1) Introduced the CNN-SVM based deep learning architecture: DNS-Tunnet and demonstrated that DNS-Tunnet outperforms all other models in terms of both performance and efficiency
2) Extensive Experiment
   - Compared Ensemble techniques with Deep Learning techniques
   - Experimented with different features to find the best combination of features for training machine learning models
   - Carried out hyperparameter tuning using Randomized and Grid Search to find the best set of values for our model's hyperparameters

978-1-6654-5872-6/22/$31.00 ©2022 IEEE

Rest of the paper is structured as follows. The previous state of art works in this domain have been discussed in Section II. Section III explains our experiment's methodology in detail. The results, observations and analysis of our proposed model have been discussed in Section IV. Finally, we present our conclusion and future research directions in Section V.

II. Related Work

Das et al. [7] proposed K means clustering, an unsupervised machine learning (ML) algorithm that solely looked at TXT records for DNS Tunneling detection, without considering other types of DNS records such as, A or AAAA. The authors clustered the labeled data into several groups and classified based on these known clusters with a comparatively lower accuracy of 91.68%. However, Preston [8] carried out a comparative analysis to show that supervised ML algorithms outperforms unsupervised ML algorithms in DNS Tunneling detection domain. Moreover Ahmed et al. [9] developed a ML model based on Isolation Forest for real-time classification between normal and anomalous DNS queries using only stateless attributes of the DNS queries that can be extracted in real-time, defined by three main categories: character count, entropy and length of separate labels in a DNS query. Although the authors were able to reach a 98.49% and 95.47% accuracy rate for University Campus and Research Institute dataset respectively, they assumed that DNS traffic is not encrypted over Transport layer Security (TLS). Wu et al. [10] further improved the accuracy and recall, both greater than 99% using logistic regression algorithm. Nadler et al. [11] even shifted focus on detecting low throughput DNS exfiltration malware using iForest algorithm and evaluated it on large-scale DNS traffic.

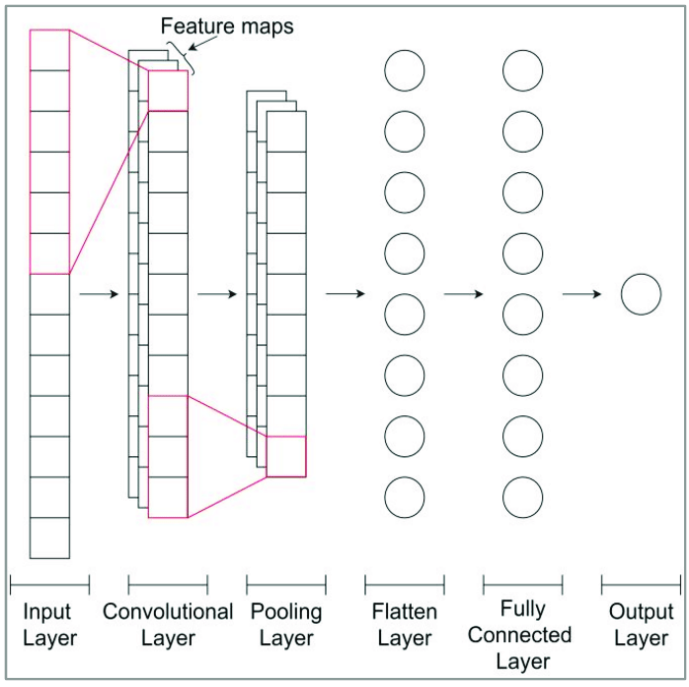


Fig. 1. 1D CNN architecture[18]

The major drawback of machine learning model was that it required tedious process of manually extracting features. To overcome this limitation, deep learning approaches have been explored in this domain. Lai et al. [12] proposed a Deep Learning model based on the feed-forward neural network (FNN) model for DNS Tunneling detection, which considers the first 512 bytes of a DNS packet as input and achieved an accuracy of 99.96%. However they focused their study based on dataset collected from only one of the DNS Tunneling tools which may affect their performance on general data. Liu et al. [13] used Convolutional Neural Network (CNN) to detect DNS tunnels where the DNS packets were represented as bytes to include overall perspective of traffic and achieved an accuracy of 99.98%. Palau et al. [14] suggested a less complex 1D CNN for DNS Tunneling detection based on lexicography. Their model was tested on tunneling and normal domains generated using 5 common tunneling tools achieving a 92% accuracy. An advanced approach to this was proposed by Lambion et al. [15] where they combined CNN and Random forest (RF), such that output of CNN classifier became one of the features of the 12-dimensional feature space vector of RF classifier. Sakarkar et al. [16] used Long-Short-Term-Memory (LSTM) in their study and achieved a high accuracy score of 99%. However the model requires high hardware requirements. Overall, deep learning models have shown to perform better than traditional machine learning models.

III. Methodology

A. Dataset

DNS Tunneling Queries Classification [17] is a dataset widely used for the detection of DNS Tunneling attacks. This dataset has labeling for two categories of domain names: regular domain names and tunneling domain names. This dataset help models classify domain names as genuine with label '0' and as tunnels for label '1'. We used an 80:20 split ratio for our training (28000) and testing data (7000).

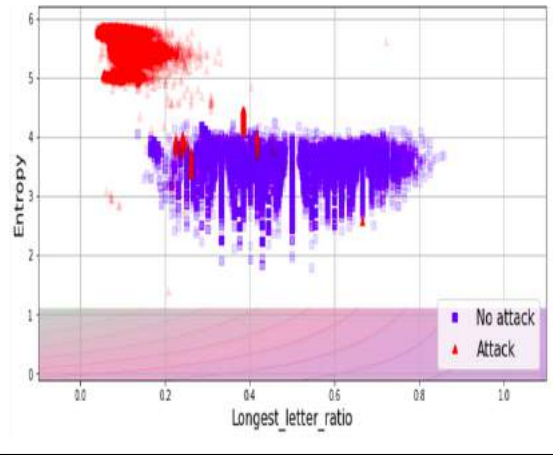
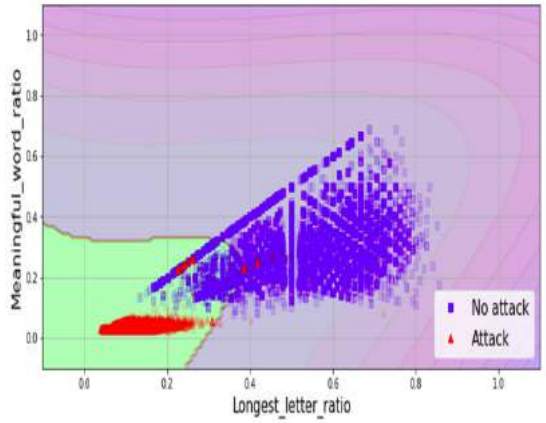
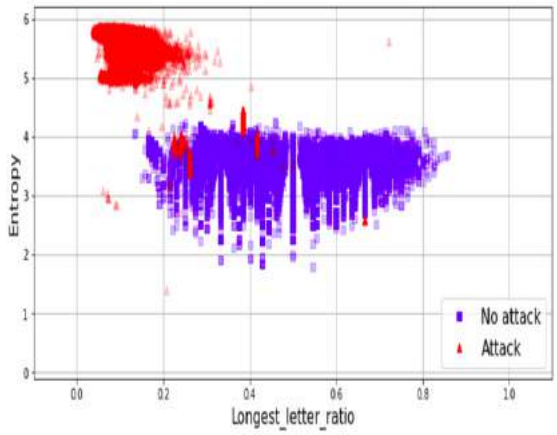
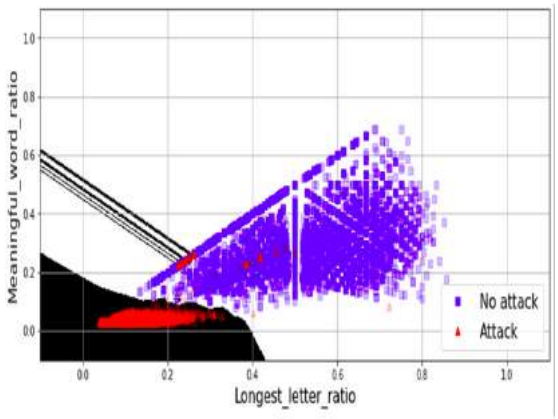
The deep learning model did not require any feature extraction process, so we used the DNS Query and Label dataset described above as it is. But for our ensemble-based models, we had to extract meaningful features from the given queries and find the correlation between these features and their corresponding labels (genuine or tunnel), as discussed later in this paper.

B. Models

1) DNS-Tunnet: DNS-Tunnet is the deep learning architecture that is proposed in this study for detection of DNS Tunneling attacks. It is a hybrid integration of CNN and SVM classifiers.

CNN: Whenever we use the word CNN, we generally refer to 2D Convolutional Neural Network, but in the present paper, we have applied 1D CNN. The architecture of a 1D CNN has been shown in Fig. 1. The shown 1D CNN architecture consists of 6 types of layer, an input layer, a convolutional layer for learning the parameters, a pooling layer for summarising the features, a flattened layer for flattening the learned features

TABLE I
2-D VISUALIZATION OF ATTACK CLASSIFICATION WITH REDUCED FEATURES

Dataset	D1	D2
SVM poly		
KNN		

to one vector, a fully connected layer and an output layer. The major difference between a 1D CNN and a 2D CNN is in the application of its kernels. Kernels can be better known as filters which are used to extract features from the dataset. In a 1D CNN, the kernels travel along one direction to extract features whereas in 2D CNN, the kernels move along two dimensions.

SVM: SVM is a supervised ML technique which is used for classification of dataset by finding a hyperplane which can be used to differentiate between the classes. SVM can be of two types : Linear SVM and Nonlinear SVM. Linear SVM uses one straight line for the classification of data whereas Nonlinear SVM are used when the data cannot be classified using a single line.

As for the regularization, it solves the problem of overfitting by nullifying the effect of parameters which increase the complexity of the model. L1 Regularization has an inbuilt feature selection property which removes the significance of the feature by multiplying its weight with 0. Whereas L2 Regularization reduces the weight of a feature and doesn't bring its value to 0. Since L2 is computationally less expensive and differentiable, we have used L2 SVM in our model.

$$\min \frac{1}{p}\|w\|_2^2 + C\sum_{i=1}^{p} \max\left(0, 1 - y_i^1\left(w^\top x_i + b\right)\right)^2 \quad (1)$$

The L2 SVM finds the parameter w by solving [19, eq.(1)] where $\|w\|_2$ is the L2 norm and C is the penalty parameter.

Hinge Loss is a loss function used for maximizing the minimum margin in an SVM model. The formula for the same:

$$l = \max\left(0, 1 - y^i\left(x^i - b\right)\right) \quad (2)$$

In [20, eq.(2)], 'l' represents the loss of instance, 'y^i' and 'x^i' represent a training set instance and b represents the bias.

2) Ensemble Models: Since this study aims to prove that deep learning techniques outperform any existing traditional machine learning techniques, the best way out for this problem is to work with one final model that takes into account the properties of all the appropriate machine learning models. For this reason, we chose to use ensemble methods as our bench mark model. Instead of working towards finding one single best model, the ensemble learning method takes into consideration different base models together to produce one optimal

model for prediction. For the purpose of this experiment we have compared two types of Ensemble methods to find the best one for further comparison with our DNS-Tunnet architecture. First is the Voting classifier that train on several models but create only a single model for predicting the output class based on the highest majority of voting. The other ensemble method we used was Random Forests. Random forest is a type of ensemble classifier which consists of multiple decision trees.

C. Experiment

1) DNS-Tunnet:

a) Preprocessing (tokenization and padding): In our first step we have carried out the tokenization process using the tokenizer function that converts each word into tokens, counts the number of each token, and ranks them accordingly. We convert every raw text to a vector of word indexes. However, since CNN requires the input sequence data with the same length, we apply padding to convert all the sequences to the same length. For our study, we have used post padding of a maximum length of 100. Keeping in mind that most of the time, the important information appears in the beginning of the query, we used the "post" padding method.

b) Implementation: Since our task involves working with textual information, we deployed a 1D-CNN. A single kernel will move one-by-one down a list of input embeddings. DNS-Tunnet comprises an embedding layer, one Convolutional layer followed by Maxpool and Flatten layers, two Dense layers, and one output layer. We have used "relu" activation function. Two drop-out layers are also added to avoid overfitting data. In our output layer, L2-SVM is implemented and the loss is computed by eq.(2). This is different from standard CNN classifiers where the conventional softmax function is implemented. Here, CNN automatically extracts meaningful features from the DNS query. This extracted vector is directly fed as an input to the SVM classifier for binary classification. This method eliminates the need for any hand-crafted data extraction process. Similar CNN-SVM integration was also studied in the image classification task on the MNIST dataset and was able to achieve approximately 99.04% of testing accuracy [21].

Fig. 2 showcases the plot for achieved training and validation accuracy and loss. Given that our validation accuracy and loss are much better compared to the training set, we can conclude that our model is not overfitted and will perform better on the newly discovered dataset. Overall training accuracy and testing accuracy came out to be 99.992% and 99.985% respectively.

2) Ensemble Methods:

a) Feature Extraction: Based on the analysis of the DNS queries, we chose some of the attributes to be used in both our datasets. We calculated the longest number string ratio, longest letter string ratio, and overall symbol ratio in the entire query. For the longest number ratio, we calculated the longest string of consecutive numbers in a string. For example in 'ab2hfb6890hdbd12' it would return '6890'. Similarly for getting the longest letter string in the DNS query, we found the longest string of consecutive letters in a string. For example in '55fg3kiotye1r2f' it would return 'kiotye'. Finally, for calculating the symbols ratio, we found all the non-alphanumeric characters in the query and calculated its ratio compared to the entire query length.

These common attributes were combined with two contrasting features discussed below to create and compare the model's performance on two different types of datasets: D1 and D2.

D1-Entropy: Entropy was first introduced as "a measure of uncertainty in a random variable" by Claude Shannon[22]. For our study, we have calculated entropy using Shannon's Entropy formula as shown in eq.(3). Here, $P(x_i)$ denotes the probability of an event x_i and $H(x)$ denotes the entropy. Usually the value of base b is taken to be 2. This equation is calculated such that higher the degree of randomness present in the string, higher will be the value of entropy and vice versa. So for example, the entropy score for domain "facebook.com", will be much lower compared to that for random domains like "aaaaxxxyydbhjfnmf24r6#6dg.ss.com".

Based on our research, numerous past studies have also shown that entropy is a great measure to detect malicious domains and subdomains formulated using a domain generator algorithm (DGA). Dataset with Entropy is referred to as D1.

$$H(x) = -\sum_{i=1}^{n} P(x_i) \log_b P(x_i) \tag{3}$$

D2-Meaningful word ratio: While entropy is a measure for randomness, we took its contrasting feature-longest meaningful word ratio for comparison of our model's performances on two different datasets. To calculate this feature, we get the longest substring that belongs to the English dictionary using the 'enchant' built-in library in python. We also checked for digits and punctuation because enchant understands digit strings and some symbols as valid words. Then we calculated the ratio of the longest meaningful word compared to the total length of the query.

b) Implementation: We built and compared two different types of ensemble models for this study: Voting Classifiers and Random Forests.

Voting Classifiers: After testing our dataset on various models we selected the top 3 performing models to feed as input to our voting classifiers. Based on our experiment, SVM, KNN, and Decision Tree were chosen for further analysis. In case of SVM, poly kernel performed the best amongst the three different types of kernels: 'linear', 'poly' and 'rbf'. Kernels convert a low-dimensional feature space data into a high-dimensional feature space dataset in order to make the dataset classifiable. The KNN algorithm is a classification algorithm in which the K nearest neighbors are selected and then classified according to the maximum number of a particular class present.

Decision tree is another ML algorithm which consists of two types of nodes: decision nodes and leaf nodes. For any

TABLE II
HYPER-PARAMETERS USED FOR ENSEMBLE MODELS

Models	Hyperparameters
SVC-poly	C=10, gamma=1, kernel='poly', random_state=13
Decision Tree	max_depth=7, min_samples_leaf=4, min_samples_split=14, random_state=13
KNN	n_neighbors=7(for D1), n_neighbors=15(for D2), weights="uniform"
Random Forest	random_state=13, max_leaf_nodes=16, n_jobs= -1, n_estimators=500

type of classification, we move on the decision nodes to reach a leaf node which makes up the final result.

For the purpose of visualization, we plotted a 2-Dimensional graph for two randomly selected features i.e. "Entropy" and "Longest_letter_ratio" for D1 and "Meaningful_word_ratio" and "Longest_letter_ratio" for D2, and plotted their classification plots achieved by SVM-poly and KNN models as shown in Table I. We can clearly see how a boundary has been identified in each case to help separate the attack set from the genuine dataset, and use it to predict for unseen cases.

We chose to train a Voting Classifier on an ensemble of different combinations of these models: Decision Tree+KNN+SVM (poly) and Decision Tree+KNN. We have used both hard and soft voting. Hard voting makes predictions on the basis of class labels and weights associated with each classifier, while soft voting relies on probabilities and weights associated with the classifiers[23].

Random Forest: For our second ensemble classifier, we chose Random Forests and trained it using both D1 and D2. The major advantage of using this classifier is that it reduces overfitting and increases accuracy. The random forest takes predictions from multiple decision trees and gives a final prediction from the one which has the majority of predictions.

c) Hyperparameter tuning: Since these models also have hyperparameters that can be self-altered, we carried out the popular Randomized and Grid search techniques to find the best parameters for our models. Hyperparameter search or optimization is a method of finding the best set of hyperparameters that enhance the performance of our model. Hyperparameters help us configure our models according to our dataset. They are different from the model's parameters that are learned automatically.

Randomized Search is carried out by defining a search space in a bounded domain for hyperparameters and random points are sampled in that domain. We used this search to discover a range of possible hyperparameter combinations. Then we carried out a Grid Search by first defining a grid search space according to the results of randomized search and then looking for the best values by evaluating every position in the grid. Randomized Search comparatively took much longer to execute than Grid search. The final hyperparameters for our model are described in Table II.

The performance of all these models is summarized in Table III.

IV. DISCUSSION

From Table III, we can see that Voting Classifier : Decision Tree + KNN performed better than Decision Tree and KNN alone. However, based on the overall comparison of our Ensemble Models, Random forests gave the best performance for both datasets. Nonetheless, our Deep Learning model, DNS-Tunnet, outperformed the ensemble-based models in terms of both performance and efficiency.

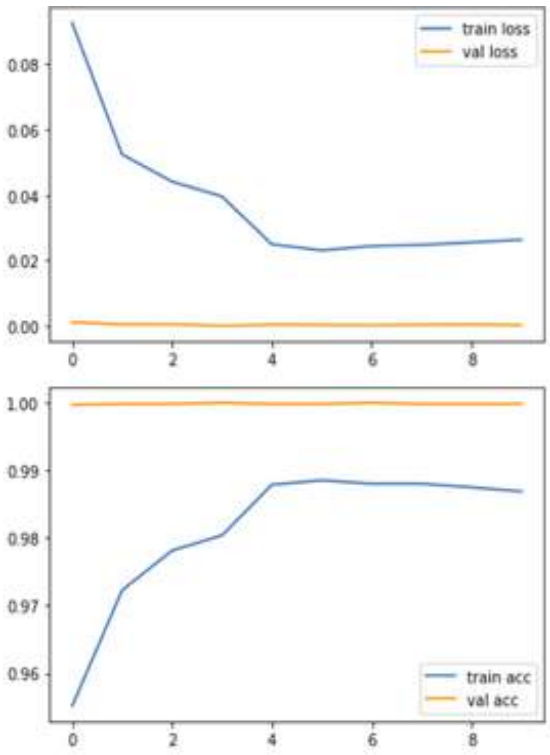


Fig. 2. DNS-Tunnet's Training loss vs Validation loss plot(top), DNS-Tunnet's Training accuracy vs Validation accuracy plot(bottom)

DNS-Tunnet has shown a testing accuracy of approximately 99.98% with an F1 score of 99.98% which is higher than that for any of the ensemble-based models. The best part is how efficient the process is, given that all the features are automatically extracted, and no background knowledge is required to carry out a manual feature engineering process as in the case of ensemble-based models.

Our data happens to be highly imbalanced in terms of positive and negative samples. So amongst all the available metrics, we have chosen to use the F1 score as our overall performance metric. Since a single DNS tunneling attack can cause immense damage and destruction, our objective is to correctly classify all the attacks and reduce the number of false negatives and positives, rather than just check how many of the observations are correctly classified.

V. CONCLUSION

In this paper, we proposed DNS-Tunnet, a CNN-SVM based architecture for DNS Tunneling attack detection. While CNN is a popular deep learning architecture, it is widely used for extraction of features. SVM on the other end is proved to be one of the best classifiers. Thus, we utilized the merits of each of these models for our proposed architecture in DNS-Tunnet.

TABLE III
EXPERIMENTAL RESULTS

Dataset	Models	Testing Accuracy %	Precision %	Recall%	AUC Score %	F1 score %
Kaggle	DNS-Tunnet	99.98	100	99.97	99.99	99.98
D1	Decision Tree (DT)	99.89	99.93	99.80	99.87	99.89
	KNN	99.67	99.35	99.90	99.69	99.63
	SVM (poly)	99.06	99.29	99.57	99.11	99.93
	SVM (linear)	98.79	97.69	99.57	98.87	98.62
	SVM (rbf)	98.90	97.91	99.61	98.98	98.76
	DT+KNN+SVM	99.67	99.38	99.87	99.69	99.91
	DT+KNN	99.73	99.54	99.84	99.74	99.69
	Random Forest	99.90	99.93	99.84	99.89	99.94
D2	Decision Tree	99.84	99.90	99.73	99.83	99.89
	KNN	99.79	99.74	99.71	99.74	99.75
	SVM (poly)	98.90	98.04	99.48	99.96	98.75
	DT+KNN+SVM	99.57	99.35	99.67	99.58	99.51
	DT+KNN	99.83	99.97	99.64	99.81	99.80
	Random Forest	99.87	99.97	99.74	99.86	99.85

Apart from this, we also built two different sets of datasets: Entropy based and Meaningful word ratio based, for our ensemble classifiers. Entropy calculates the amount of randomness present while meaning word ratio calculates the ratio of longest meaningful word present in the query. Our experiments showed that the dataset consisting of entropy achieved better results. Amongst the Voting and Random Forests Ensemble classifiers, Random Forests performed better. Overall, DNS-Tunnet outperformed all the existing models by giving a F-score of 99.98%.

In future, we aim to self collect data from numerous tunneling tools to help train our model on broader category of tunneling attack types. Moreover, currently our model only carries out binary classification for tunneled and genuine DNS queries. Nonetheless in future we also aim to work towards incorporating classification of the type of DNS Tunneling present if any, like FTP, HTTP, HTTPS or POP3.

REFERENCES

[1] A. Lal, A. Prasad, A. Kumar, and S. Kumar, "Data Exfiltration: Preventive and Detective Countermeasures."

[2] "What is DNS Tunneling? — DDI (Secure DNS, DHCP, IPAM) — Infoblox." https://www.infoblox.com/glossary/dns-tunneling/ (accessed Mar. 18, 2022).

[3] "72% of Organizations Experienced a DNS Attack in the Past Year - Infosecurity Magazine." https://www.infosecurity-magazine.com/news/72-orgs-dns-attack-last-year/ (accessed Mar. 18, 2022).

[4] M. Lyu and H. H. Gharakheili, "Mapping an Enterprise Network by Analyzing DNS Traffic," vol. 2.

[5] J. Ahmed, H. H. Gharakheili, Q. Raza, C. Russell, and V. Sivaraman, "Monitoring Enterprise DNS Queries for Detecting Data Exfiltration from Internal Hosts," pp. 1–14, 2017.

[6] M. Lyu, H. H. Gharakheili, C. Russell, and V. Sivaraman, "Hierarchical Anomaly-Based Detection of Distributed DNS Attacks on Enterprise Networks," vol. 18, no. 1, pp. 1031–1048, 2021.

[7] A. Das, M. Y. Shen, M. Shashanka, and J. Wang, "Detection of exfiltration and tunneling over DNS," Proc. - 16th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2017, vol. 2017-Decem, pp. 737–742, 2017, doi: 10.1109/ICMLA.2017.00-71.

[8] R. Preston, "DNS Tunneling Detection with Supervised Learning," vol. 7013, no. 19, pp. 1–6, 2019.

[9] J. Ahmed, H. H. Gharakheili, Q. Raza, C. Russell, and V. Sivaraman, "Real-Time Detection of DNS Exfiltration and Tunneling from Enterprise Networks," 2019.

[10] K. Wu, Y. Zhang and T. Yin, "CLR: A Classification of DNS Tunnel Based on Logistic Regression," 2019 IEEE 38th International Performance Computing and Communications Conference (IPCCC) , 2019, pp. 1-1, doi: 10.1109/IPCCC47392.2019.8958731.

[11] A. Nadler, A. Aminov, and A. Shabtai, "Detection of malicious and low throughput data exfiltration over the DNS protocol," Comput. Secur., vol. 80, no. 2, pp. 36–53, 2019, doi: 10.1016/j.cose.2018.09.006.

[12] C. Lai, B. Huang, S. Huang, C. Mao, and H. Lee, "Detection of DNS Tunneling by Feature-free Mechanism," 2018 IEEE Conf. Dependable Secur. Comput., pp. 1–2, 2018.

[13] C. Liu, L. Dai, W. Cui, and T. Lin, "A Byte-level CNN Method to Detect DNS Tunnels," 2019 IEEE 38th Int. Perform. Comput. Commun. Conf., pp. 1–8.

[14] F. Palau, C. Catania, J. Guerra, S. Garcia, and M. Rigaki, "DNS Tunneling: A Deep Learning based Lexicographical Detection Approach," 2020, [Online]. Available: http://arxiv.org/abs/2006.06122.

[15] D. Lambion, M. Josten, F. Olumofin, and M. De Cock, "Malicious DNS Tunneling Detection in Real-Traffic DNS Data," Proc. - 2020 IEEE Int. Conf. Big Data, Big Data 2020, pp. 5736–5738, 2020, doi: 10.1109/BigData50022.2020.9378418.

[16] G. Sakarkar, M. K. H. Kolekar, G. Patil, P. Dutta, R. Chaturvedi, and S. Kumar, "Advance Approach for Detection of DNS Tunneling Attack from Network Packets Using Deep Learning Algorithms," vol. 10, pp. 241–266, 2021.

[17] "DNS Tunneling Queries Classification — Kaggle." https://www.kaggle.com/saurabhshahane/dns-tunneling-queries-classification (accessed Mar. 18, 2022).

[18] Nisa, E. C. and Kuan, Y. Der (2021) 'Comparative assessment to predict and forecast water-cooled chiller power consumption using machine learning and deep learning algorithms', Sustainability (Switzerland), 13(2), pp. 1–18. doi: 10.3390/su13020744.

[19] A. F. M. Agarap, "An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification," pp. 5–8.

[20] "A definitive explanation to the Hinge Loss for Support Vector Machines. — by Vagif Aliyev — Towards Data Science." https://towardsdatascience.com/a-definitive-explanation-to-hinge-loss-for-support-vector-machines-ab6d8d3178f1 (accessed Mar. 18, 2022).

[21] "MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges." http://yann.lecun.com/exdb/mnist/ (accessed Mar. 18, 2022).

[22] S. Vajapeyam, "Understanding Shannon ' s Entropy metric for Information," no. March, pp. 1–6, 2014.

[23] "Hard vs Soft Voting Classifier Python Example - Data Analytics." https://vitalflux.com/hard-vs-soft-voting-classifier-python-example/ (accessed Mar. 18, 2022).

View publication stats

Domain-Specific Hybrid BERT based System for Automatic Short Answer Grading

Jai Garg
Department of Applied Mathematics
Delhi Technological University Delhi, India
jaigarg_ 2k18mc044@dtu.ac.in

Jatin Papreja
Department of Applied Mathematics
Delhi Technological University Delhi, India
jatinpapreja_ 2k18mc049@dtu.ac.in

Kumar Apurva
Department of Applied Mathematics
Delhi Technological University Delhi, India
kumarapurva_2k18mc058@dtu.ac.in

Dr. Goonjan Jain
Department of Applied Mathematics
Delhi Technological University Delhi, India
goonjanjain@dtu.ac.in

***Abstract*—Effective and efficient grading has been recognized as an important issue in any educational institution. In this study, a grading system involving BERT for Automatic Short Answer Grading (ASAG) is proposed. A BERT Regressor model is fine-tuned using a domain-specific ASAG dataset to achieve a baseline performance. In order to improve the final grading performance, an effective strategy is proposed involving careful integration of BERT Regressor model with Semantic Text Similarity. A set of experiments is conducted to test the performance of the proposed method. Two performance metrics namely: Pearson's Correlation Coefficient and Root Mean Squared Error are used for evaluation purposes. The results obtained highlights the usefulness of proposed system for domain specific ASAG tasks in real life.**

***Index Terms*—Automatic Short Answer Grading (ASAG), Semantic Text Similarity, Key-Response Similarity, Bidirectional Encoder Representation from Transformers (BERT), Masked and Permuted Pre-training for Language Understanding (MPNet)**

I. INTRODUCTION

Grading assignments and tests are an important part of any educational course. It is used as a method for assessing the level of understanding developed by an individual undertaking a particular course. No grading system is perfect but effective automation can provide numerous benefits on a large scale.

Generally, an assessment involves questions that can be classified into three categories on the basis of dependency on the reference key:

- *High Dependency*: involves questions such as MCQs, Fill-Ups and True/False.
- *Moderate Dependency*: involves questions having short answers generally from one-liners to a paragraph.
- *Low Dependency*: involves language based questions such as letter writing, essay writing and debate writing.

Modern Examinations generally involves Multiple Choice Questions due to ease of grading. In these questions, candidates are required to select one option out of some given options. It has already been shown that these types of questions are inadequate in accessing the caliber of students due to their closed ended nature [1].

Questions involving short answers can be a suitable replacement for the MCQ's if they can be graded efficiently. Short answers typically refer to one or two line responses given to a question in natural language involving free text. Short answer type questions generally require some additional information (key) along with the question for grading properly. Also, the response can be graded based on the context, making it a subjective decision rather than an objective one.

Traditionally, manual grading has been a preferred choice for assessment of student's responses. However, with increase in the number of students pursuing educational stream, workload of teachers and professors has increased many folds. Sometimes, manual grading may also introduce some bias or inconsistency due to various reasons such as involvement of more than one grader, writing style of a student and lack of appreciation of answers different from the reference key. Also, students are left deprived of accurate and timely feedback on their tests and assignments. Thus, a solution to this problem should involve an automated grading system that can provide faster results with relatively low level of bias and inconsistency while also enabling students to showcase their skills and knowledge. Automatic Short Answer Grading (ASAG) can be used to overcome these challenges. It involves a system to learn and understand high-level contextual meaning of reference key as well as a student's response for a particular question. Thus, it can be classified as Natural Language Processing task.

In this paper, the problem of Automatic Short Answer Grading is tackled using Bidirectional Encoder Representations from Transformers (BERT) on a domain-specific ASAG dataset. Mohler dataset [2] is used as ASAG dataset which contains questions and answers involving Data Structures.

BERT [3] is a transformer-based machine learning technique for Natural Language Processing (NLP) developed and pre-

This work is licensed under a Creative Commons Attribution-Noncommercial 4.0 International License

trained by Google. Unlabeled data extracted from the books with 800 million words and Wikipedia with 2,500 million words is used for pre-training. The original English-language BERT has two models:

1) *BERT Base*: 12 layers (transformer blocks), 12 attention heads, and 110 million parameters.
2) *BERT Large*: 24 layers (transformer blocks), 16 attention heads and, 340 million parameters.

It is pretrained on two tasks namely Masked Language Modeling (MLM) and Next Sentence Prediction (NSP).

- *Masked Language Modeling (MLM)*: BERT is designed as a deeply bidirectional model. The network effectively captures information from both the right and left context of a token starting from the first layer and all the way through to the last layer.
- *Next Sentence Prediction (NSP)*: BERT is also trained on the task of Next Sentence Prediction for tasks that require an understanding of the relationship between sentences.

As a result of the training process, BERT learns contextual embeddings for words. After pretraining, it can be finetuned with fewer resources on smaller datasets to optimize its performance on specific tasks.

The contributions of this paper are summarized as follows:

- Building an Automatic Short Answer Grading System involving a baseline fine-tuned BERT Regressor model.
- Generating Semantic Text Similarity using a modified BERT model fine-tuned for Question Answering and a Sentence Transformer model.
- Integrating BERT Regressor model with Semantic Text Similarity to further improve the performance.
- Finally, creating a model which is capable of grading a student's response for a particular question when provided with a reference key.

Thus, the system should be able to learn and understand high-level contextual meaning of both reference key as well as student response and compare them to generate a final grade in a specified numerical range.

The paper is further organized as follows: Section II contains the related work. Section III includes the approach proposed for the ASAG task. Experimental Setup including Dataset, Architecture and System Specifications is discussed in Section IV. Section V contains the results obtained from the proposed method and its comparison with different existing methods, followed by conclusion in Section VI.

II. RELATED WORK

Studies on Automatic grading has gained a lot of momentum in recent years. The need to have a reliable ASAG system is at its peak right now. The studies have been fueled further by the onset of pandemic and shift of teaching practices to online mode. Authors from around the world have proposed various techniques to solve the ASAG task. Developing an effective as well as a reliable ASAG system remains the focal center in the studies being carried out in recent years.

The study of automatic grading system was initiated by Page [4] for grading essays using computers. Since then, various contribution in this domain have been made by many authors around the world. Burrows et al. [5] researched about the existing ASAG system and classified them into 5 categories based on timeline and the techniques used. They researched about 35 ASAG systems from 1996 to 2015 which were the flag bearers of development in the field of automatic grading. Mohler et al. [2] proposed an ASAG system based on lexical semantic similarity measures. They showed that a combination with machine learning techniques is more useful in accurately predicting grades as compared to that in isolation. They also studied dependency graphs of response and key and further researched over their alignment to gain more information.

Recently, Ramachandran et al. [6] proposed a word-order graph-based study to achieve the ASAG task. They rely on finding important patterns from rubric texts and responses from highly graded students. They also explored semantic metrics to find out the synonyms to represent replacement words. Sultan et al. [7] proposed an ASAG system based on feature extraction. The features extracted included Text similarity between reference answer and student response, question demoting, term weighting and length ratio. Based on these features, regression and classification models were built according to the requirement of dataset. Wang et al. [8] introduced ml-BERT method for ASAG task. They combined BERT with meta-learning to help in initialization of the BERT parameters in a specific target subject domain using unlabeled data, thus leveraging the limited labeled training data for the grading task. Sung et al. [9] studied about improving BERT by supplementing data from resources belonging to a particular domain for ASAG task. They used multi-domain resources as datasets to perform fine-tuning. Tulu et al. [10] presented an ASAG system using sense vectors obtained from SemSpace algorithm and LSTM combined with Manhattan Vectorial Similarity. Sense embeddings of Synsets corresponding to each word in Student's answers or reference answers are given as input into parallel LSTM architecture. Finally, Manhattan Similarity is found between the text embedding of student's response and reference answer.

With the help of these prior studies, a hybrid approach consisting of fine-tuned BERT Regressor model followed by integration of key-response semantic text similarity for a particular question is proposed.

III. METHODOLOGY

The proposed ASAG system is responsible for generating a numerical grade (G) when provided with a Question (Q), a Key (K) and a Response (R).

$$(Question, Key, Response) \rightarrow Grade \quad (1)$$

The proposed system involves a hybrid approach consisting of three fundamental steps as shown in Fig. 1. In the first step (A), BERT Regressor model is fine-tuned using a domain specific ASAG dataset. It involves a BERT model followed by

a regressor model. Second step (B) involves determining the similarity between key and response for a particular question. A modified Question-Answering BERT model along with cosine similarity is proposed for this purpose. In the third step (C), the similarity characteristic obtained in second step is combined with BERT Regressor model to provide final grading results. This section provides a detailed explanation of the three mentioned steps.

A. BERT Regressor ASAG Model

A BERT base model is combined with a custom artificial neural network based regression model for the ASAG task. The architecture of BERT Regressor model is discussed in Section IV. The model is trained using a domain-specific ASAG dataset to generate the grades. In this study, objective is to generate numerical grading in a particular range. Thus, a regression model is used.

Initially, a sequence of tokens is generated for each key-response pair using BERT tokenizer. This sequence includes a classification token [CLS] at the start of sequence and a separator token [SEP] in between key and response as shown in Fig. 2. This sequence of tokens is then passed on as an input to the BERT Regressor model for training purposes. The training task involves predicting the grade in a particular range for an input sequence of tokens.

The BERT model generates a high dimensional embedding of each input sequence which is further fed to the regressor model in order to generate a grade for the key-response pair.

B. Determining Key-Response Similarity

The semantic similarity between a key and response is one of the most important features on the basis of which grading should be performed. More the semantic similarity between key and response for a particular question, higher should be the numerical grade awarded.

Evaluating similarity between a key-response pair is an important task and involves caution. A response might contain keywords present in the key, but can still lack contextual similarity. Another case may arise, where the length of response is relatively longer than the reference key but semantically correct. Thus, simple truncation of a response from an extreme end would result in an inaccurate representation of original response. In order to overcome these challenges, a filtered response is constructed from original response which is then used for finding similarity with the key.

In order to construct the filtered response, a fine-tuned BERT model for Question-Answering is used. A typical Question-Answering BERT model works by generating a start and stop value for each word in the response. The model selects the word with maximum start value as the starting word of the answer. Similarly, word with maximum stop value is chosen as the ending word while ensuring that the starting word comes before the ending word. This method of extracting answers from the response can lead to very short filtered responses and thus, result in a loss of information as compared

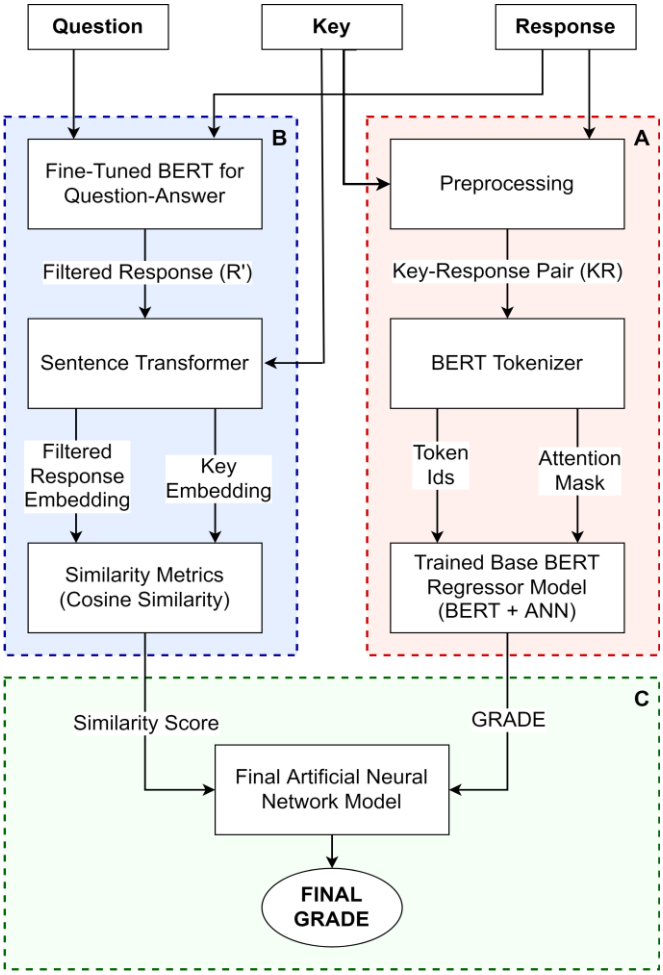


Fig. 1. Model Approach Flowchart.

Question: What is the experimental approach for measuring the running time of an algorithm?

Key: Implement the algorithm and measure the physical running time.

Student Response: running an algorithm on a specific set of data

Key-Response Pair: [CLS] Implement the algorithm and measure the physical running time. [SEP] running an algorithm on a specific set of data. [SEP]

Token IDs of Key-Response Pair without padding:
[101, 10408, 1996, 9896, 1998, 5468, 1996, 3558, 2770, 2051, 1012, 102, 2770, 2019, 9896, 2006, 1037, 3563, 2275, 1997, 2951, 102]

Fig. 2. Example of Key-Response Pair and Token ID's.

to original response. Eventually, this can result in a lower grade for even a very good response.

To overcome this problem of potential information loss, a modified technique is proposed for selecting starting and ending words. In the proposed method, starting and ending words are chosen in such a manner that length of the answer (filtered response) obtained is at least as long as the key provided. Hence, the generated filtered response contains majority of the important information required for grading purposes while eliminating any irrelevant text as shown in Fig. 3.

Question: What is the difference between a circular linked list and a basic linked list?

Key: The last element in a circular linked list points to the head of the list.

Student Response: The circular linked list's tail points to the head, whereas the basic linked list's tail points to a NULL.

Filtered Response: tail points to the head, whereas the basic linked list's tail points to a null

Key Embeddings:
[0.01993977 -0.1645347 -0.01536680 ... -0.03414393 0.01878243 -0.04168871]

Filtered Response Embeddings:
[0.01826572 -0.1475896 0.012160487 ... -0.04462539 0.04133997 -0.01269365]

Cosine Similarity Score: 0.723192

Fig. 3. Example of Filtered Response and Similarity Score.

The filtered response generated contains only the important section extracted out from a response and thus, ensures that quality is preferred over quantity. This also helps in partially achieving the task of question demoting. Filtered response and its corresponding key is fed to the Masked and Permuted Pre-training for Language Understanding (MPNet) [11] model to generate two embeddings respectively. These embeddings are then combined using a similarity metric to generate a similarity score. In the proposed approach, cosine similarity is used as similarity metric, which represents the cosine of the angle between the two embeddings when drawn in the corresponding embedding space. Equation (2) is used to calculate the cosine similarity where *A* and *B* represents the two embeddings.

$$Similarity\ Score\ (A, B) = \frac{A \cdot B}{|A||B|} \quad (2)$$

Overall, this method is used to generate a similarity score between the important information extracted from the response and the key.

C. Integration of Similarity Score with BERT Regressor

The similarity score is integrated with the BERT Regressor model using an artificial neural network. The network is then trained in order to predict the final grade. Any score outside the permitted range is rounded off to the nearest score within the range. Using similarity score as a separate feature, results in performance improvement of the proposed system as discussed in Section V.

IV. EXPERIMENTAL SETUP

A. Dataset

The proposed system is evaluated using Mohler ASAG dataset [2]. It involves generating a numerical score in range 0 to 5 for a student's response to a particular question on the basis of a reference key provided. The dataset consists of 80 questions from ten assignments and two exams based on Data Structures. It contains 2,273 student responses. The average grade of two different human graders has been considered as the final grade. The entire dataset has been divided into training and testing set with 0.2 being the test size.

B. Model Architecture

In BERT Regressor, BERT-base-uncased model is used for generation of embeddings. It generates 768-dimensional embedding vector for each input sequence of tokens. Artificial neural network-based regression model contains an input layer, 2 hidden layers, an output layer along with dropout layers to prevent overfitting as shown in Fig. 4. Rectified Linear Unit (ReLU) is used as the activation function.

Key-Response similarity method uses BERT-large-uncased model involving whole word masking and finetuning on SQuAD dataset as the Question-Answering BERT model. MPNet model is used in order to generate embeddings for filtered response and key before the application of cosine similarity. Neural network used for integration of similarity score with BERT Regressor model contains 2 hidden layers with ReLU activation function and an output layer with a linear activation function as shown in Fig. 5.

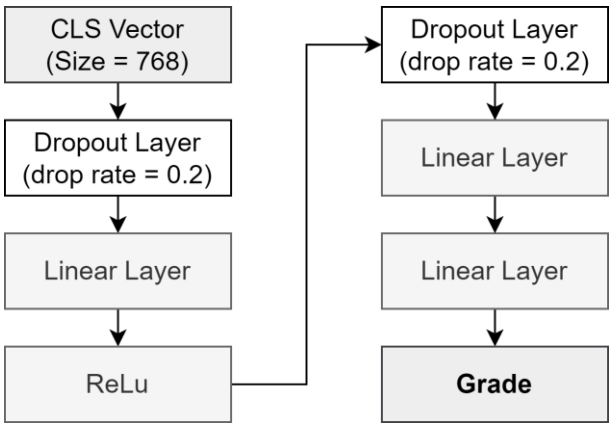


Fig. 4. BERT Regressor ANN architecture.

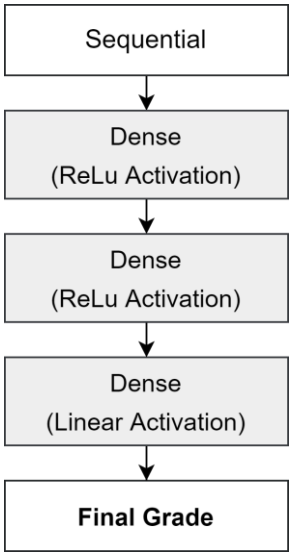


Fig. 5. ANN architecture for integration phase.

C. Parameters

BERT Regressor model is trained using AdamW (Adam with weight decay) optimizer with a learning rate of $5x10^{-5}$ and epsilon being 10^{-8}. Number of epochs is chosen to be 10 with a batch size of 32. Mean Squared Error (MSE) is used as the loss function accompanied by a linear scheduler with warmup. All other parameters retain their default values.

D. System Specifications

Google Colab is used for developing the proposed model with 2 vCPUs, GPU Tesla K80 having compute capability of 3.7 and 12 GB RAM.

V. RESULTS

A. Comparative Study

The proposed model as well as base model (BERT Regressor) is compared to the existing systems based on its performance on the test set. Pearson's Correlation Coefficient (R) and Root Mean Squared Error (RMSE) are used as two evaluation metrics. These are well known metrics for evaluation of any ASAG system as determined by various other studies. A higher Pearson's Correlation Coefficient and a lower RMSE value is desired for a good ASAG system. The results obtained using proposed approach along with other existing models are shown in Table 1. The proposed model gives Pearson's Correlation Coefficient as 0.777 and RMSE value as 0.732 on the test set.

Base model gives 0.760 as Pearson's Correlation Coefficient and 0.753 as RMSE. A clear improvement (about 2-3%) over the base model is observed in both evaluation metrics upon integration with the similarity score.

Hence, similarity score results in a more robust grading due to better understanding of the key-response pair while reducing the biasness of base model towards higher scores.

TABLE I
COMPARISON OF PROPOSED SYSTEM WITH EXISTING ASAG SYSTEMS

Models	**Pearson's R**	**RMSE**
Final Proposed Model (BERT Regressor + Similarity Score)	0.777	0.732
Base Model (BERT Regressor)	0.760	0.753
Tulu et al. (2021) [10]	0.949	0.040
Sultan et al. (2016) [7]	0.630	0.850
Tf-Idf (2016) [7]	0.320	1.020
Ramachandran et al. (2015) [6]	0.610	0.860
Mohler et al. (2011) [2]	0.518	0.978

The Proposed model performs better than most of the existing models except MaLSTM model proposed by Tulu et al [10]. The reason being that their LSTM model is trained on each assignment separately rather than on the entire dataset in one go. Thus, it narrows down the domain to each sub-topic. The difference in results between the proposed model and MaLSTM model can be justified since the proposed model is developed keeping domain generality as the primary objective.

B. Sample Data Analysis

In order to provide a more comprehensive overview of the results obtained, a random sample of six student responses are extracted from the test set. Fig. 6 shows the actual and the predicted grades of the random sample. Predicted scores almost match actual scores in majority of the questions. The RMSE for the random sample of questions comes out to be 0.32. Fig. 7 shows the result of a response to a particular question along with a key. The score given through manual grading is 5.0 and the score given by the proposed model comes out to be 4.979. This result validates the semantic similarity present between the response and the reference key.

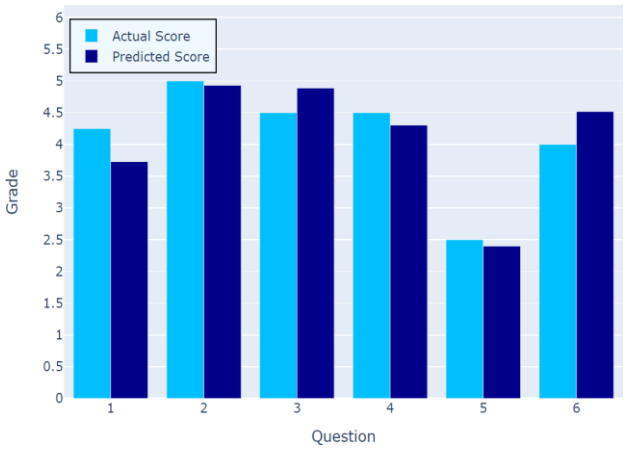


Fig. 6. Actual v/s Predicted Grade Comparison for a sample set.

Question: What does a function signature include?

Key: The name of the function and the types of the parameters.

Student Response: the function's name and parameters

Score Given By Manual Grading: 5.0

Score Given By the proposed approach: 4.9794846

Fig. 7. Example of Grading using the proposed system.

VI. CONCLUSION AND FUTURE WORK

In this paper, a 'Domain-Specific Hybrid BERT based System for Automatic Short Answer Grading' is developed. The system relies on understanding high-level contextual meaning of Question, Key and Response to perform automatic grading. It involves integration of similarity score with a base BERT Regressor model. The performance of the model is analysed using Pearson's Correlation Coefficient and Root

Mean Squared Error. Mohler's ASAG Dataset is used for experimental analysis which is widely known for benchmarking ASAG systems. The proposed model performs better than the genuinely eligible models considered for comparison by providing a higher value of Pearson's Correlation Coefficient and a lower value of RMSE. The results obtained using the proposed model highlights the usefulness of the approach in real life.

As a future study, the performance can be further improved by selecting more suitable and performance centric models with more computational power for the purpose of Question Answering, Sentence Similarity and BERT-Regressor model. Experimenting with different similarity metrics and parameter tuning of proposed models may also lead to performance gains. Implementation of classification version of proposed method for datasets such as SemEval [12] can be useful.

REFERENCES

[1] Y. Oksuz and E. Demir, "Comparison of open ended questions and multiple choice tests in terms of psychometric features and student performance," Hacettepe Univ. J. Edu., vol. 34, no. 1, pp. 259–282, 2019, doi: 10.16986/HUJE.2018040550.

[2] M. Mohler, R. Bunescu, and R. Mihalcea, "Learning to grade short answer questions using semantic similarity measures and dependency graph alignments," Proc. 49th Annu. Meeting Assoc. Comput. Linguistics, Hum. Lang. Technol., 2011, pp. 752– 762.

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, arXiv:1810.04805. [Online]. Available: http://arxiv.org/abs/1810.04805

[4] E. B. Page, "The imminence of grading essays by computers," Phi Delta Kappan, vol. 47, no. 5, pp. 238–243, 1966.

[5] S. Burrows, I. Gurevych, and B. Stein, "The eras and trends of automatic short answer grading," Int. J. Artif. Intell. Edu., vol. 25, no. 1, pp. 60–117, Mar. 2015, doi: 10.1007/s40593-014-0026-8.

[6] L. Ramachandran, J. Cheng, and P. Foltz, "Identifying patterns for short answer scoring using graph-based lexico-semantic text matching," Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications, 2015.

[7] M. A. Sultan, C. Salazar, and T. Sumner, "Fast and easy short answer grading with high accuracy," in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Human Lang. Technol., 2016, pp. 1070–1075, doi: 10.18653/v1/N16-1123.

[8] W. Zichao, S. L. Andrew, E. W. Andrew, P. Grimaldi, and R. G. Baraniuk, "A meta-learning augmented bidirectional transformer model for automatic short answer grading," in Proc. 12th Int. Conf. Educ. Data Mining (EDM), 2019, pp. 1–4.

[9] C. Sung, T. Dhamecha, S. Saha, T. Ma, V. Reddy, and R. Arora, "Pretraining BERT on domain resources for short answer grading," in Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP), 2019, pp. 6073–6077, doi: 10.18653/v1/D19-1628.

[10] C. N. Tulu, O. Ozkaya and U. Orhan, "Automatic Short Answer Grading With SemSpace Sense Vectors and MaLSTM," in IEEE Access, vol. 9, pp. 19270-19280, 2021, doi: 10.1109/ACCESS.2021.3054346.

[11] Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie-Yan Liu, "MPNet: Masked and Permuted Pre-training for Language Understanding", Proceedings of NeurIPS (2020), pp. 16857-16867.

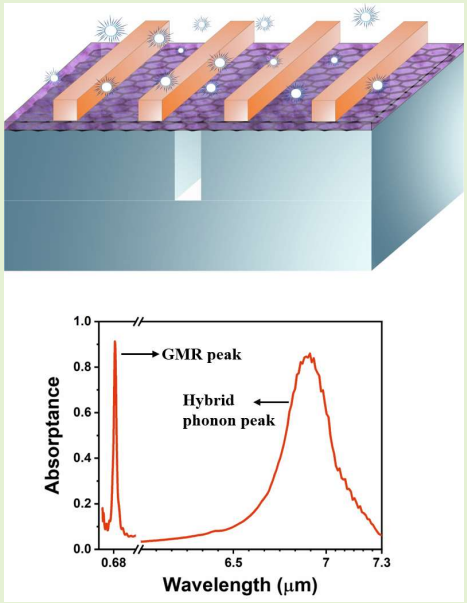
[12] R. Nielsen, M. Dzikovska, C. Brew, C. Leacock, D. Giampiccolo, L. Bentivogli, P. Clark, I. Dagan, and H. Dang, "SemEval-2013 Task 7: The joint student response analysis and 8th recognizing textual entailment challenge," Assoc. Comput. Linguistics, Atlanta, GA, USA, Tech. Rep., 2013, pp. 263–274.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

Dual-Band Biosensing with van der Waals Assisted Optical Metasurfaces

Rashmi Kumari, Shubhanshi Sharma, Shailendra Kumar Varshney, and Basudev Lahiri

***Abstract*— In this paper, we theoretically investigate a van der Waals (vdW) heterostructure-based optical metasurface for dual-band optical biosensing capable of attaining refractive index (RI) sensing in the visible and vibrational fingerprint detection of the same analyte in the mid-infrared (mid-IR) region simultaneously. The proposed structure consists of two gratings – Calcium Fluoride (CaF_2) grating to produce guided-mode resonance in visible and silver (Ag) grating coupled with graphene/ hexagonal boron nitride (hBN) heterostructure to induce hybrid plasmon phonon mode in mid-IR that can be used for surface-enhanced infrared absorption (SEIRA) spectroscopy. The evanescent wave in CaF_2 gratings gives rise to RI sensing at ~ 689 nm. The presence of graphene adds tunability and thus allows identification of multiple vibrational fingerprints of organic molecules, nitroethane and nitrobenzene in mid-IR, with a sensitivity of 360.6 nm/RIU and 313.4 nm/RIU, respectively, in the visible region. The fingerprint signal gets enhanced at 6.43 μm (nitroethane) and 6.56 μm (nitrobenzene) by varying the Fermi potential of the graphene layer above 0.3 eV. The proposed sensor has the potential to meet the requirement of next-generation medical applications such as disease diagnostics, health monitoring, and food/environment monitoring as the sensor enables label-free, sensitive, rapid detection of organic molecules.**



***Index Terms*— Dual-band, Graphene, Refractive index sensing, surface-enhanced infrared absorption (SEIRA) spectroscopy, vibrational fingerprints, tunability, van der Waals heterostructure.**

I. Introduction

Optical sensing is a powerful technology for highly sensitive, non-destructive, reliable, and label-free detection of organic molecules and monitoring food/health/environment[1], [2]. The recent research and technological developments have enabled rapid and accurate detection of organic nanoparticles at a low cost. Optical metasurfaces that consist of artificial photonic components of subwavelength dimensions is a powerful device for advancing sensing technologies. These optical metasurfaces can control light flow into nanoscale volumes for rapid, quantitative, and real-time results for biomedical applications [3]. Plasmonic metasurfaces using surface plasmon resonance (SPR)-collective oscillation of electrons at the surface of metals have shown good results in refractive index sensing [4], [5]. SPR-biosensors have been reported for the enhanced detection of a variety of biomarkers, including drugs, bacteria, and proteins, by functionalizing the sensor with specific aptamers [6]–[8].

(*Corresponding author: Rashmi Kumari, Basudev Lahiri).*
Rashmi kumari and Basudev Lahiri is with Nano Bio Photonics Group,
Shubhanshi Sharma and Shailendra Kumar Varshney is with Fiber Optics, Nano & Quantum Photonics (FONQP) Group,
Advance Photonics Laboratory, Department of Electronics and Electrical Communication Engineering, IIT Kharagpur, West Bengal, 721302, India

Localized surface plasmon resonance (LSPR) is a further development of SPR, where surface plasmons oscillate locally around the metallic nanoparticles whose size is less than the wavelength of incident light [9]. The highly localized electromagnetic fields at the nanoparticle surfaces enable improved detection of organic molecules [10]. The LSPR sensors are simple, cost-effective, easy to integrate into miniaturized devices and show high sensitivity towards change in local refractive index that has been shown to detect many biomarkers [11], [12]. However, the refractive index sensing suffers from poor selectivity of the target molecules. Different molecules could have similar refractive indexes, and different concentrations of the same molecules can result in variation of a shift in SPR wavelength. Therefore, in contrast to refractive index sensing, infrared spectroscopy enables the detection of characteristic molecular vibrations acting as fingerprints of specific molecules in the mid-infrared (mid-IR) region [13], [14]. Interaction of analyte with the mid-IR signals on optical metasurface permits analyte-specific detection of molecular vibrations. Thus, optical metasurfaces can guide light at the nanoscale, whose evanescent fields can be used to detect the shift in the resonance wavelength and also the interaction with an analyte in mid-IR to probe the molecular structure of the targeted analyte.

However, the organic molecules like protein, carbohydrates,

XXXX-XXXX © XXXX IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

and enzymes that are the building blocks of life have low weight, possess small absorption, and interact weakly with the incident light hence are less sensitive to detection [15]. To improve the sensitivity, plasmonic nanostructures are used to absorb the incident light energy to nanoscale volumes. This highly enhanced electromagnetic energy manifests the absorption of the molecular vibration when placed in such high confinement region. This highly sensitive detection of molecular vibrations using enhanced absorption of electromagnetic fields is called surface-enhanced infrared absorption (SEIRA) spectroscopy [16]–[20]. Organic molecules possess their vibrational fingerprints in the mid-IR region, and therefore, SEIRA-based sensing is an effective technique to probe into the molecular structure of the target analyte [21], [22]. To further enhance the sensitivity of the plasmonic metasurfaces, recently developed two-dimensional (2D) materials such as graphene with highly tunable optical and electrical properties are employed to increase the light-matter interaction over a wide range of wavelengths [2], [23]. Refractive index sensing using dual-band resonance produced by graphene nanostructure was done for detecting multiple analytes simply by varying the Fermi potential of the graphene layer [24]. Another 2D material, hexagonal boron nitride (hBN)- a natural hyperbolic van der Waals (vdW) material, possesses anisotropic phonon polaritons (PhPs) in the mid-IR region [25]. These anisotropic PhPs provide deep subwavelength confinement, optical super-resolution, and ultrasensitive infrared spectroscopy in the mid-IR range [26], [27]. Graphene, in combination with hBN, creates a vdW heterostructure that couples PhPs in the hBN layer with graphene plasmons and generates hybrid plasmon phonon polaritons, which exhibit electrical conductivity tunability along with subwavelength field confinement [28], [29]. Such hybrid plasmon phonon polariton has been previously utilized for ultrasensitive detection of multiple analytes in mid-IR [30].

Therefore, optical metasurfaces possess the ability to identify the analyte via changes in the resonance wavelength in the visible region and can probe their vibrational fingerprints in the mid-IR. Thus, it becomes important to simultaneously perform refractive index sensing and selective identification of molecular vibrational fingerprints on a single substrate. Various works have been done on this idea to achieve resonance wavelength shift along with infrared spectroscopy using a single substrate [31][32]. Angle tunability was also employed to detect multiple vibrational peaks in mid-IR [33]. However, there is a need for a dual-band optical biosensor in which one resonance can provide refractive index sensing in the visible region while the other resonance can be used for vibrational fingerprint detection in the mid-IR regime.

In this work, we propose a vdW assisted optical metasurface with hBN/graphene heterostructure sandwiched between CaF_2 grating and silver grating to show guided-mode resonance (GMR) in the visible region due to CaF_2 grating and the magnetic plasmon (MP) resonance in mid-IR due to silver grating. The refractive index of any material defines its behavior towards the incident electromagnetic waves. Thus, the measurement of change in the refractive index indicates the presence of an external perturbation. Numerical simulations reveal a significant shift in GMR wavelength in response to the change in the surrounding refractive index. In comparison, the coupling of MP resonance in silver grating with the phonons in the hBN layer placed on top of it is responsible for the infrared spectroscopy in the mid-IR regions [30], [33]. Incorporating graphene with hBN in the structure makes it a vdW heterostructure that adds the required tunability to the device. Thus, we can simultaneously identify multiple vibrational fingerprints in a broad range of mid-IR wavelengths along with their refractive index sensing in the visible region. The comparison of the functionality of our proposed sensor with some other works are given in table I.

TABLE I
Comparison of the proposed sensor with previous works

Ref	Refractive Index Sensing	SEIRA	Tunability
[24]	Yes (mid-IR)	No	Wavelength Tunability
[30]	No	Yes (mid-IR)	Wavelength Tunability
[31]	Yes (mid-IR)	Yes (mid-IR)	No
[32]	Yes (mid-IR)	Yes (mid-IR)	No
[33]	Yes (mid-IR)	Yes (mid-IR)	Angle Tunability
Our Work	Yes (Visible)	Yes (mid-IR)	Wavelength Tunability

II. BIOSENSOR DESIGN AND NUMERICAL MODELING

The schematic of the proposed biosensor is shown in Fig. 1(a), which consists hBN/graphene heterostructure sandwiched between silver grating and CaF_2 grating. The geometrical parameters of the proposed sensor are optimized in a way such that the two gratings produce its response in two different regions without interfering with each other. The period of the two-subwavelength grating ($a < \lambda_{GMR}$, $p < \lambda_{PhP}$) is chosen such that it prevents higher diffraction orders from propagation [34]. The 2D front view of the design is shown in Fig 1(b), where silver grating of height h_1 = 1.5 μm is placed on the silver substrate at the bottom with periodicity p = 2.5 μm and split gap b = 0.1 μm. The hBN of thickness d =20 nm is placed above the grating, and a graphene sheet is sandwiched between hBN and silver grating. On the top of hBN/graphene heterostructure, CaF_2 grating is made with periodicity a = 625 nm, width w = 120 nm, and thickness t = 150 nm. Subwavelength silver grating produces MP resonance in mid-IR, whereas subwavelength CaF_2 grating generates GMR in the visible regime.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

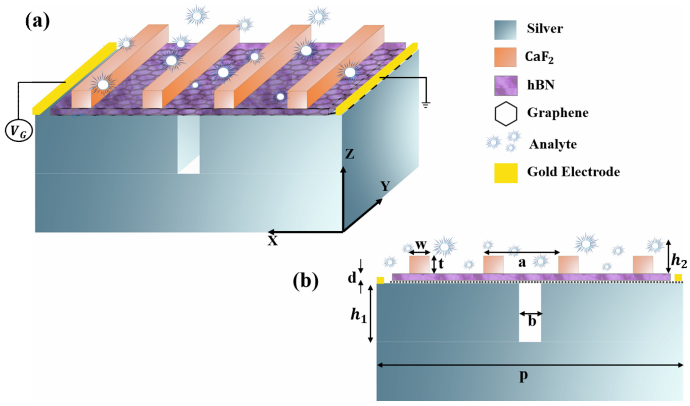


Fig. 1. (a) Schematic of the proposed biosensor with hBN/ graphene vdW heterostructure sandwiched between CaF2 grating on top and silver grating on the bottom. (b) 2D front view of the design with the mentioned geometrical parameters.

A. hBN permittivity analysis

hBN is a naturally occurring anisotropic van der Waals material whose in-plane and out-of-plane permittivity are opposite in signs in the region called as Reststrahlen band (RB), producing phonon polaritons in the band [35]. Two RBs exist in hBN - Type I (12 μm - 12.82 μm) and Type II (6.2 μm -7.3 μm). Here, we have focused on Type II RB, as it strongly confines the electromagnetic waves (PhPs) within its volume that propagate with very large momentum, possess sub-diffraction confinement, and have been utilized in super-resolution optical imaging [25], [26]. The dielectric permittivity of hBN is modeled analytically as the Lorentz oscillator model [36]:

$$\varepsilon_{\xi}=\varepsilon_{\infty,\xi}\left(1+\frac{\omega_{LO,\xi}^{2}-\omega_{TO,\xi}^{2}}{\omega_{TO,\xi}^{2}-i\gamma\omega-\omega^{2}}\right) \quad (1)$$

And is charaterized by a diagonal tensor as:

$$\varepsilon_{\xi}=\begin{pmatrix}\varepsilon_{\perp} & 0 & 0\\ 0 & \varepsilon_{\perp} & 0\\ 0 & 0 & \varepsilon_{\parallel}\end{pmatrix} \quad (2)$$

where, ξ can be either ⊥ (in-plane component), or ∥ (out-of-plane component), accordingly, $\omega_{TO,\perp}$ = 1370 cm^{-1}, $\omega_{LO,\perp}$ = 1610 cm^{-1}, $\omega_{TO,\parallel}$ = 780 cm^{-1}, $\omega_{LO,\parallel}$ = 830 cm^{-1}, $\varepsilon_{\infty,\perp}$ = 4.87, $\gamma_{\perp}$= 5 cm^{-1}, $\varepsilon_{\infty,\parallel}$ = 2.95, and $\gamma_{\parallel}$ = 4 cm^{-1}.

B. Graphene Conductivity Analysis

Graphene possesses excellent electrical and optical properties. Here, we use surface conductivity approach to model single layer graphene. The optical conductivity of graphene, given by Kubo's formula, depends on both intraband and interband conductivity, as mentioned below [28]:

$$\sigma_g=\sigma_{\text{intra}}+\sigma_{\text{inter}}$$

$$=\frac{2e^{2}k_{B}T}{\pi\hbar^{2}}\frac{i}{(\omega+i/\tau)}\ln\left|2\cosh\left(\frac{E_f}{2k_{B}T}\right)\right|+\ldots$$
$$\frac{e^{2}}{4\hbar}\left[H(\omega/2,T)+\frac{4i}{\pi}\int_{0}^{\infty}d\zeta\,\frac{H(\zeta,T)-H(\omega/2,T)}{\omega^{2}-4\zeta^{2}}\right] \quad (2)$$

where

$$H(\omega,T)=\sinh(\hbar\omega/k_{B}T)\Big/\left[\cosh(E_f/k_{B}T)+\cosh(\hbar\omega/k_{B}T)\right],$$

E_f is the Fermi energy, k_B is the Boltzman constant, $\hbar$ is the planck's constant, e is the electron charge, ω is the incident light angular frequency, $\tau=\mu E_f/(e\,v_f)^2$ is the electron relaxation time, where Fermi velocity $v_f = 10^6$ m/s and carrier mobility μ=10^4 cm^2/ (V.s) at temperature T=300k. Gate bias voltage (V_G) is applied across the graphene sheet and silver separately that controls the Fermi level of graphene given as [37], [38]:

$$V_G=\left(eE_f^{2}t\Big/\pi\hbar^{2}v_f^{2}\varepsilon_0\varepsilon_{air}\right) \quad (3)$$

where ε_{air} is the permittivity of air separation between graphene and silver, t is the separation between the graphene and silver. It is assumed that the thickness of the air layer between graphene and silver is 1 nm. Therefore, biasing voltage corresponding to E_f=0 eV and E_f=0.5 eV are V_G = 0 V and V_G = 1.66 V, respectively.

To show the working mechanism of the proposed device as a sensor in both visible and mid-IR regions, we consider nitroethane (n=1.33) and nitrobenzene (n=1.52) as two organic molecules with their fingerprints at 6.43 μm and 6.56 μm respectively in mid-IR. We have used the finite element method (FEM) based numerical simulations to calculate transmission spectrum and field distribution. In FEM simulations, light is incident in the z-direction with its electric field polarized along the x-direction. Periodic boundary conditions are set in x and y directions, and a perfectly matched layer is set in z directions to absorb all the scattered radiations.

III. Simulation Results and Discussions

A. Sensing in visible region

When the electric field is in the x-direction, the sub-wavelength grating structure produces guided-mode resonance that can reflect or transmit specific electromagnetic fields [39], [40]. These guided modes cannot be excited by the normal incidence of plane waves [41]. Due to the sub-wavelength grating structure, the diffraction produced by the grating is in-phase with the guided mode of the waveguide, hence guiding a GMR mode in the grating [42]. Here we have used CaF_2 grating that does not have its resonance in the visible or mid-IR region and is optimized to produce GMR at 689 nm wavelength in the visible regime, as shown in Fig. 2(a) [43]. The substrate being silver, GMR is absorptive in nature at the resonance wavelength with its electric field confined mainly in the vicinity of CaF_2, as shown in Fig. 2(b). Another resonance at 866 nm (Fig. 2(a)) is due to silver grating, whose electric

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

and magnetic fields are not guided, as shown in Fig. 2(c) and therefore is not considered for sensing.

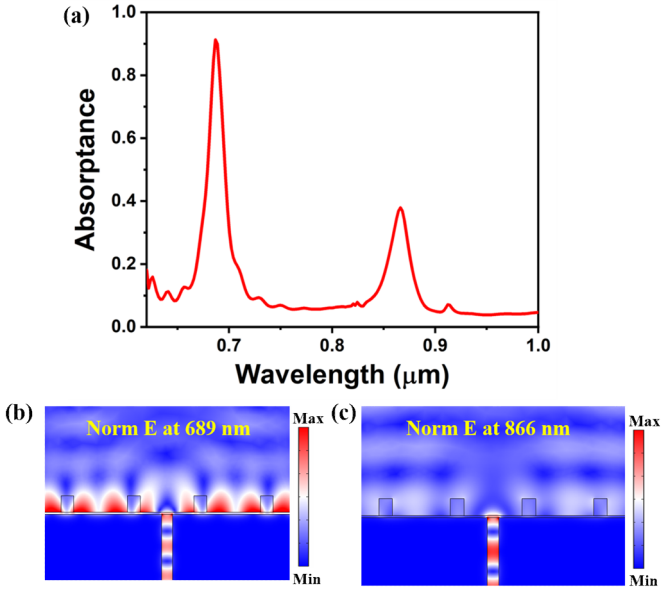


Fig. 2. (a) Absorptance spectrum, dictating GMR peak at 689 nm. The normalized electric field at (b) GMR at 689 nm and (c) 866 nm.

It can be seen that the normalized electric field is mainly confined at the side-ends of the CaF_2 grating. The evanescent fields outside the structure are responsible for detecting the refractive index changes in the surrounding. The evanescent fields interact with the analytes directly and are very sensitive to the changing refractive index in the surrounding medium induced by changing analytes concentration. This changing refractive index results in shift of resonance wavelength, which is expressed in terms of the sensitivity defined as [44], [45]:

$$S=\frac{\Delta\lambda}{\Delta n} \quad (4)$$

where $\Delta\lambda$ is the resonance wavelength shift corresponding to the change in refractive index Δn. The thickness of the analyte layer is taken as h_2 =180 nm. The greater the evanescent field component penetration into the air, the larger the sensitivity.

Refractive index sensing of two different analytes, nitroethane (n=1.33) and nitrobenzene (n=1.52) in the visible region, is shown in Fig. 3. A red shift in the absorption peak is observed with a sensitivity of 360.6 nm/RIU and 313.4 nm/RIU for nitroethane and nitrobenzene, respectively. The absorption decreases as the refractive index of the analyte increase, *viz.* nitrobenzene. Note that graphene/hBN vdW heterostructure is sandwiched between CaF_2 grating and the silver grating. The GMR mode penetrates feebly into the 20 nm thin hBN layer, slightly reducing the evanescent field [46]. The presence of graphene does not affect the sensor's performance, as in the visible region, graphene's interband absorption is dominant where the resonance wavelength is not affected by varying Fermi potential [47]. Though refractometric sensing can detect the presence of analytes via the change in its resonance wavelength, it suffers from poor selectivity. It cannot differentiate between various molecules possessing similar refractive indices. To make these nanophotonic sensors applicable for practical applications, we need to achieve both refractive index sensing in visible and infrared spectroscopy in the mid-IR to know the complete information of analyte.

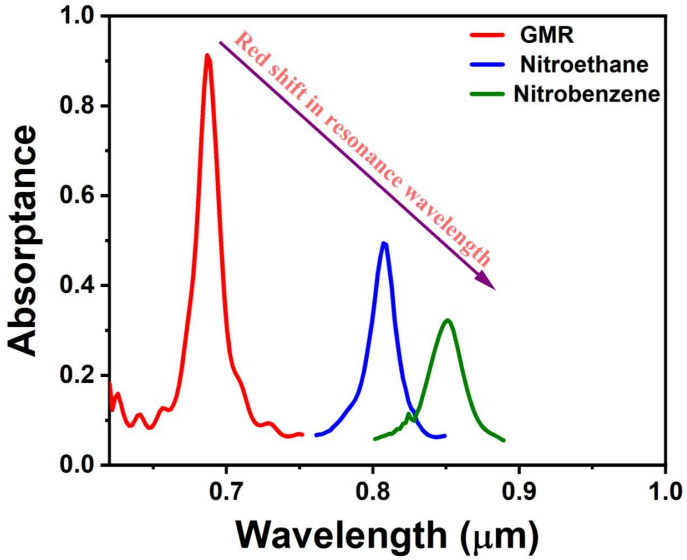


Fig. 3. The absorptance spectrum portrays refractive index sensing in the visible region for two different analytes - Nitroethane (n=1.33) and Nitrobenzene (n=1.52).

B. Sensing in mid infrared region

Organic molecules exhibit their material-specific vibrational modes, particularly in the mid-IR spectrum. The proposed sensor not only detects the binding of analytes in the visible region, as discussed earlier but also probes into their vibrational fingerprints to reveal respective structural information using mid-IR spectroscopy. The interaction between nanometer-sized molecules and mid-IR wavelengths is weak, which leads to low sensitivity [48]. To enhance this interaction, we use an absorber surface on which molecules are allowed to adsorb. This approach to enhancing a molecule's vibrational absorption by placing it near the strong electromagnetic confinement of the absorber is called surface-enhanced infrared absorption spectroscopy (SEIRA) [21], [49].

To perform SEIRA spectroscopy, x polarized plane-wave is incident normally on the proposed device, which leads to magnetic plasmon (MP) resonance in the silver grating. A 20nm thick hBN layer exhibits its phonon resonance at 7.2 μm, as shown in Fig. 4 by the dotted green curve [50], whereas silver grating alone exhibits MP resonance at 7.9 μm, shown by the dashed blue curve[51]. The hBN layer, when it rests on a silver grating structure, induces a coupling between the phonon and MP resonance [52]. As the electric field confinement is strong at the corner of the grating, its momentum is sufficient to excite phonon polaritons in hBN and generate two hybrid modes with near-perfect absorption, as portrayed in Fig. 4 by solid red curve. The hybrid phonon

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

mode at 6.9 μm lies inside the RB of hBN and is characterized by a zigzag pattern of normalized electric field corresponding to phonon polaritons. These PhPs propagate inside hBN film and gradually decays as shown in Fig. 5(a) [53]. This extended light-matter interaction over a large surface of hBN is useful in optical sensing [54], whereas hybrid plasmon mode at 8.6 μm stays outside the RB, with its normalized electric field concentrated near the grating as shown in Fig. 5(b).

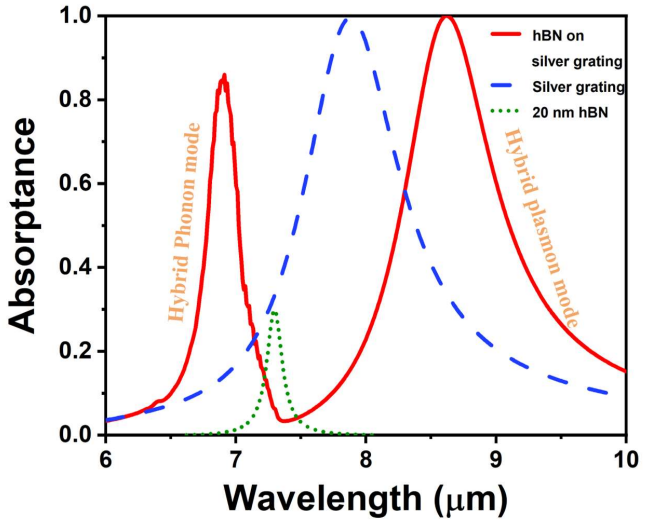


Fig. 4. Absorptance spectrum of the sensor in mid-IR (red curve) showing hybrid phonon mode and hybrid plasmon mode, 20 nm hBN (dotted green curve), only silver grating (dashed blue curve).

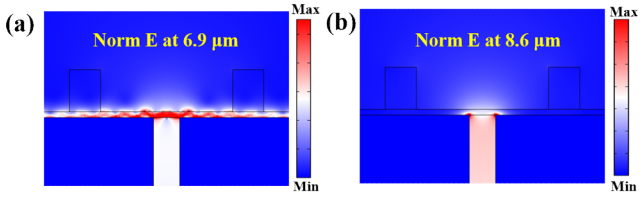


Fig. 5. Normalized electric field distribution at (a) 6.9 μm due to hybrid phonon mode confined in hBN and (c) 8.6 μm due to hybrid MP mode confined at the corner of grating

To tune both the resonance modes (6.9 μm and 8.6 μm) in the absorptance spectrum, a graphene layer is placed beneath hBN layer, creating vdW heterostructure [55]. Graphene plasmons coupled with the hybrid plasmon-phonon mode create a route to tuning the resonance peak. With the increase in the Fermi potential E_f of graphene, the resonant modes experience blue-shift [30]. The interested hybrid phonon mode lies inside the RB of hBN. Due to subwavelength confinement and low loss propagation of phonon polaritons, it is possible to enhance the optical sensing of organic molecules even at very low concentrations. The coupled hybrid phonon modes at 6.9 μm in graphene/hBN vdW heterostructure possess both the electrical tunability of graphene plasmons and long-decay time of hBN phonons and thus are used here for fingerprint detection[56]. The blue shift in resonance wavelength with the increasing E_f is attributed to the 'power-law scaling,' which states that the carrier concentration (η_c) increases by raising E_f, which in principle enhances the graphene plasmon resonance (ω_p), i.e., $\omega_p \alpha |E_f|^{1/2} \alpha \eta_c^{1/4}$ [57]. Therefore, it is possible to actively tailor the resonance mode over a wide wavelength range by varying the E_f of graphene between 0.1 eV–0.5 eV, adding key functionality to detect multiple vibrational modes without changing the structural parameters, as shown in Fig. 6. The absorption peak is blue-shifted with the rise in potential applied to the graphene layer, characterized by Fermi potential.

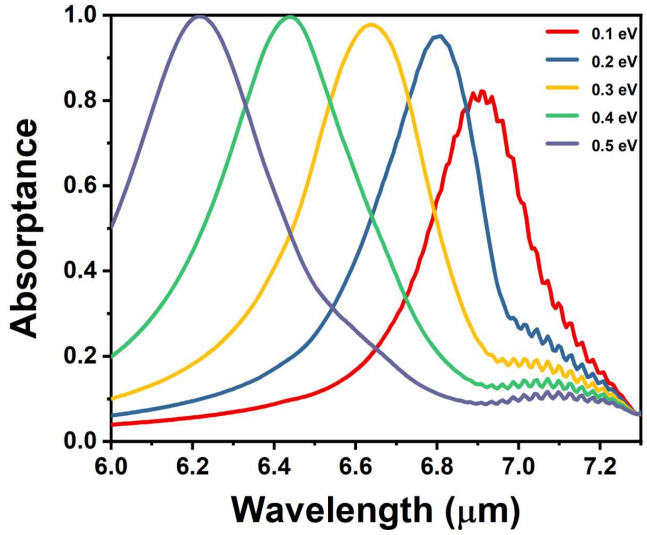


Fig. 6. Tunable absorptance spectra of hybrid phonon mode for various graphene Fermi potentials, E_f = 0.1 eV - 0.5 eV.

This feature makes the proposed metasurface an excellent SEIRA substrate that can selectively detect vibrational bonds of multiple analytes. To demonstrate the capability of the designed biosensor, we consider the same organic molecules used for RI sensing, *i.e.,* nitroethane and nitrobenzene. Nitroethane is an organic compound whose vapors may cause irritation to skin, eyes, and mucous membranes, with its characteristic fingerprint at 6.43 μm. Nitrobenzene is also an organic compound that emits toxic fumes that may cause skin, eyes, respiratory tract irritation, and shortness of breath with its characteristic fingerprint of 6.56 μm. These two analytes are hazardous, so their detection is essential. The real and imaginary part of the refractive index of these organic molecules is shown in the inset of Fig. 7 [58]. As these organic molecules are lightweight, they exhibit very low absorption of the incident light, as shown in Fig. 7. In order to detect these molecules effectively, we place them on our proposed SEIRA substrate, where high absorption of hybrid phonon mode is actively tuned to overlap with the specific molecular vibrations of the analytes in the mid-IR. Analyte molecules interact with the highly confined electromagnetic field in the absorber, hence manifest their absorption on top of the enhanced line shape of an absorber. Thus, we can detect multiple molecules with enhanced vibrational absorption. We have shown the tuning of hybrid phonon mode with different graphene's Fermi potential to detect the vibrational fingerprint of nitroethane at 6.43 μm and nitrobenzene at 6.56 μm as shown in Fig. 8. We have summarized the results for both the analytes in table II.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

TABLE II
Performance analysis of sensor in both regions

Parameters	Nitroethane	Nitrobenzene
Refractive index in visible	1.33	1.52
Wavelength of vibrational fingerprint in mid-IR	6.43 μm	6.56 μm
Sensitivity	360.6 nm/RIU	313.4 nm/RIU
E_f required	Above 0.3 eV	Above 0.3 eV

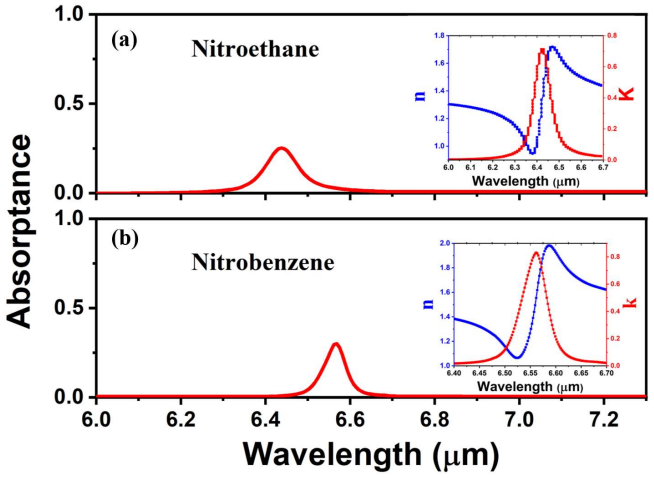


Fig.7. Absorptance spectra of (a) Nitroethane and (b) Nitrobenzene showing small absorption intensity. Inset shows real and imaginary part of refractive index for Nitroethane and Nitrobenzene

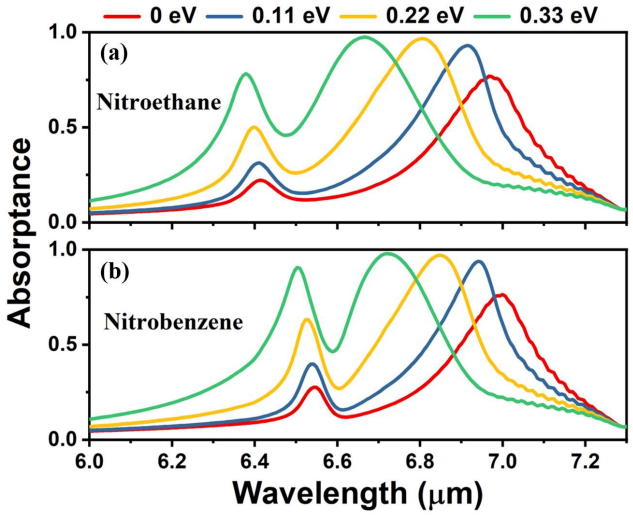


Fig. 8. SEIRA spectroscopy for two organic molecules (a) Nitroethane with its vibrational fingerprint at 6.43 μm (b) Nitrobenzene with its vibrational fingerprint at 6.56 μm for different values of graphene's Fermi potential.

We have studied the effect of parameters – width of CaF_2 (w) and silver split gap (b)- on the device's sensitivity to calculate fabrication tolerance. Changing these geometrical parameters slightly affects the sensitivity of the device in the visible region. Whereas, at mid-IR, the change in geometrical parameters does not affect the detection of vibrational fingerprints of analytes. We have simulated the device with periodic boundary conditions in x and y directions and perfectly matched layers (PML) in z directions to absorb all the scattered radiations from the device. By keeping all parameters constant and varying w and b, we observe a slight shift in sensitivity, as shown in Fig. 9. There is a slight change in sensitivity with the change of 0.3 μm in w and 0.05 μm in b. To ensure thermal management in our device, we have considered the thermal properties of the materials used. Silver, CaF_2, hBN, and graphene are stable at room and high temperatures [59]–[61]. Graphene has negative thermal expansion coefficients and excellent tolerance to high temperatures up to 1200 K [62]. We have shown the increase in conductivity of the graphene layer with increasing temperature in Fig. 10(a) and the corresponding device's response in Fig. 10(b)[63]. These results prove that our device can operate at high temperatures with excellent thermal stability.

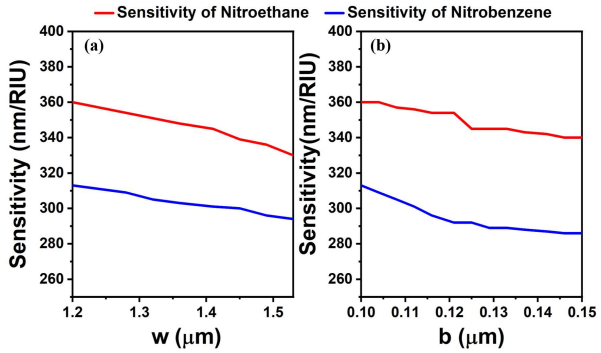


Fig. 9. Effect of variation of (a) w and (b) b on the sensitivity of the proposed device

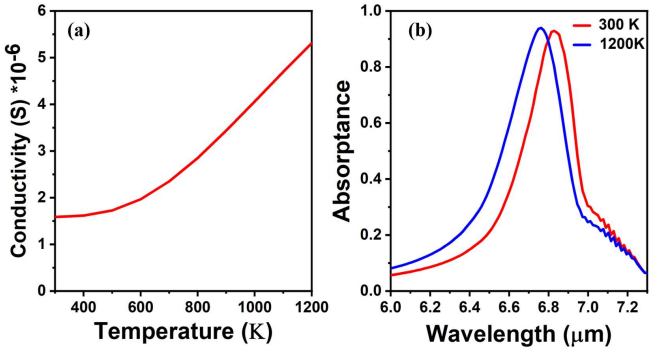


Fig. 10 (a) Temperature dependent conductivity (b) Absorptance spectra in mid-IR of the proposed device with the variation in temperature.

IV. CONCLUSION

To summarize the work, optical metasurface made of CaF_2 sub-wavelength grating placed on hBN/graphene assisted silver grating is designed numerically to detect two different organic molecules. The proposed structure is suitable to be used both for visible refractive index sensing and mid-IR spectroscopy. Two absorption spectra in two different bands are attributed to the GMR by CaF_2 grating in the visible region and hybrid phonon resonance mode in hBN on the silver

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

grating in the mid-IR region. The presence of the graphene layer adds an important tunability functionality that not only enhances the resonance peaks but also renders a wide wavelength window. Numerical simulations reveal the enhanced detection of vibrational fingerprints of nitroethane at 6.43 μm and nitrobenzene at 6.56 μm with a sensitivity of 360.6 nm/RIU and 313.4 nm/RIU, respectively.

REFERENCES

[1] A. Leitis, M. L. Tseng, A. John-Herpin, Y. S. Kivshar, and H. Altug, "Wafer-Scale Functional Metasurfaces for Mid-Infrared Photonics and Biosensing," *Adv. Mater.*, vol. 33, no. 43, 2102232, 2021

[2] S. H. Oh, H Altug, X Jin, T Low, SJ Koestee, AP Ivanov, JB Edel, P Avouris, MS Strano, "Nanophotonic biosensors harnessing van der Waals materials," *Nat. Commun.*, vol. 12, no. 1, pp. 1–18, 2021

[3] S. Zhang, CL Wong, S Zeng, R Bi, K Tai, K Dholakia, M Olivo, "Metasurfaces for biomedical applications: Imaging and sensing from a nanophotonics perspective," *Nanophotonics*, vol. 10, no. 1, pp. 265–299, 2021

[4] B. Lahiri, S. G. Mcmeekin, A. Z. Khokhar, R. M. De La Rue, and N. P. Johnson, "Asymmetric split ring resonators for optical sensing of organic materials," *Opt. Express*, vol. 17, no. 2, pp. 1107–1115, 2009

[5] Z. Sadeghi and H. Shirkani, "Highly sensitive mid-infrared SPR biosensor for a wide range of biomolecules and biological cells based on graphene-gold grating," *Phys. E Low-Dimensional Syst. Nanostructures*, vol. 119, 114005, 2020

[6] A. K. Sharma and C. Marques, "Design and performance perspectives on fiber optic sensors with plasmonic nanostructures and gratings: A review," *IEEE Sens. J.*, vol. 19, no. 17, pp. 7168–7178, 2019

[7] M. Lobry, M Loyez, K Chah, EM Hassan, E Goormaghtigh, MC DeRosa, R Wattiez, C Caucheteur, "HER2 biosensing through SPR-envelope tracking in plasmonic optical fiber gratings," *Biomed. Opt. Express*, vol. 11, no. 9, p. 4862, 2020

[8] M. Loyez. EM Hassan, E.M, M. Lobry, F Liu, C Caucheteur, R. Wattiez, MC DeRosa, WG Willmore, J. Albert, "Rapid Detection of Circulating Breast Cancer Cells Using a Multiresonant Optical Fiber Aptasensor with Plasmonic Amplification," *ACS Sensors*, vol. 5, no. 2, pp. 454–463, 2020

[9] M. Li, R. Singh, C. Marques, B. Zhang, and S. Kumar, "2D material assisted SMF-MCF-MMF-SMF based LSPR sensor for creatinine detection," *Opt. Express*, vol. 29, no. 23, p. 38150, 2021

[10] W. P. Hall, S. N. Ngatia, and R. P. Van Duyne, "LSPR biosensor signal enhancement using nanoparticle-antibody conjugates," *J. Phys. Chem. C*, vol. 115, no. 5, pp. 1410–1414, 2011

[11] M. Arques, B. I. Z. Hang, and S. A. K. Umar, "Convex fiber-tapered seven core fiber-convex fiber (CTC) structure-based biosensor for creatinine detection in aquaculture," vol. 30, no. 8, pp. 13898–13914, 2022.

[12] Y. Wang,G. Zhu, M.Li, R. Singh, C. Marques, R. Min, BK Kaushik, B. Zhang, R. Jha, S Kumar., "Water Pollutants p-Cresol Detection Based on Au-ZnO Nanoparticles Modified Tapered Optical Fiber," *IEEE Trans. Nanobioscience*, vol. 20, no. 3, pp. 377–384, 2021

[13] R. Adato and H. Altug, "In-situ ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas," *Nat. Commun.*, vol. 4, 2013

[14] C. Wu,AB Khanikaev, R. Adato, N Arju, AA Yanik, H. Altug, G. Shvets, "Fano-resonant asymmetric metamaterials for ultrasensitive spectroscopy and identification of molecular monolayers," *Nat. Mater.*, vol. 11, no. 1, pp. 69–75, 2012

[15] K. Ataka, T. Kottke, and J. Heberle, "Thinner, smaller, faster: IR techniques to probe the functionality of biological and biomimetic systems," *Angew. Chemie - Int. Ed.*, vol. 49, no. 32, pp. 5416–5424, 2010

[16] L. Kühner, R. Semenyshyn, M. Hentschel, F. Neubrech, C. Tarín, and H. Giessen, "Vibrational Sensing Using Infrared Nanoantennas: Toward the Noninvasive Quantitation of Physiological Levels of Glucose and Fructose," *ACS Sensors*, vol. 4, no. 8, pp. 1973–1979, 2019

[17] L. V. Brown, X. Yang, K. Zhao, B. Y. Zheng, P. Nordlander, and N. J. Halas, "Fan-shaped gold nanoantennas above reflective substrates for surface-enhanced infrared absorption (SEIRA)," *Nano Lett.*, vol. 15, no. 2, pp. 1272–1280, 2015

[18] J. Chae, B. Lahiri, J. Kohoutek, G. Holland, H. Lezec, and A. Centrone, "Metal-dielectric-metal resonators with deep subwavelength dielectric layers increase the near-field SEIRA enhancement," *Opt. Express*, vol. 23, no. 20, p. 25912, 2015

[19] J. Chae, B. Lahiri, and A. Centrone, "Engineering Near-Field SEIRA Enhancements in Plasmonic Resonators," *ACS Photonics*, vol. 3, no. 1, pp. 87–95, 2016

[20] I. G. Mbomson, S. Tabor, B Lahiri, G. Sharp, SG MsMeekin, R.M. Da La Rue, NP. Johnson, "Asymmetric split H-shape nanoantennas for molecular sensing," *Biomed. Opt. Express*, vol. 8, no. 1, p. 395, 2017

[21] I. Hwang, M. Kim, J Yu, J. Lee, JH Choi, SA Park, WS Chang, J Lee, JY Jung, "Ultrasensitive Molecule Detection Based on Infrared Metamaterial Absorber with Vertical Nanogap," *Small Methods*, vol. 5, no. 8,2100277, 2021

[22] D. Etezadi, J. B. Warner, F. S. Ruggeri, G. Dietler, H. A. Lashuel, and H. Altug, "Nanoplasmonic mid-infrared biosensor for in vitro protein secondary structure detection," *Light Sci. Appl.*, vol. 6, no. 8, pp. 17029-10, 2017

[23] D. Rodrigo, O. Limaj, D. Janner, D Etezadi, FJ Garcia de Abajo, V. Pruneri, H. Altug, "Mid-infrared plasmonic biosensing with graphene," *Science (80-.).*, vol. 349, no. 6244, pp. 165–168, 2015

[24] S. Gong, B. Xiao, L. Xiao, S. Tong, S. Xiao, and X. Wang, "Hybridization-induced dual-band tunable graphene metamaterials for sensing," *Opt. Mater. Express*, vol. 9, no. 1, p. 35, 2019

[25] J. D. Caldwell, I. Aharonovich, G. Cassabois, J. H. Edgar, B. Gil, and D. N. Basov, "Photonics with hexagonal boron nitride," *Nat. Rev. Mater.*, vol. 4, no. 8, pp. 552–567, 2019

[26] S. Dai, Q Ma, T. Andersen, AS Mcleod, Z Fei, MK Liu, M Wagner, K Watanabe, T. Taniguchi, M. Thiemens, F. Keilmann, "Subdiffractional focusing and guiding of polaritonic rays in a natural hyperbolic material," *Nat. Commun.*, vol. 6, no. 1, pp. 1–7, 2015

[27] J. Jiang, X. Lin, and B. Zhang, "Broadband negative refraction of highly squeezed hyperbolic graphene plasmons," *Resear ch*, pp. 1–14, 2018.

[28] A. Kumar, T. Low, K. H. Fung, P. Avouris, and N. X. Fang, "Tunable light-matter interaction and the role of hyperbolicity in graphene-hbn system," *Nano Letters*, vol. 15, no. 5. pp. 3172–3180, 2015

[29] S. Dai, Q Ma, MK Liu, T. Andersen, Z. Fei, MD Wagner, K Watanabe, T Taniguchi, M. Thiemens, F. Keilmann, "Graphene on hexagonal boron nitride as a tunable hyperbolic metamaterial," *Nat. Nanotechnol.*, vol. 10, no. 8, pp. 682–686, 2015

[30] R. Kumari, A. Yadav, S. Sharma, T. Das Gupta, S. K. Varshney, and B. Lahiri, "Tunable Van der Waal's optical metasurfaces (VOMs) for biosensing of multiple analytes," *Opt. Express*, vol. 29, no. 16, p. 25800, 2021

[31] W. C. Shih, G. M. Santos, F. Zhao, O. Zenasni, and M. M. P. Arnob, "Simultaneous chemical and refractive index sensing in the 1-2.5 μm near-infrared wavelength range on nanoporous gold disks," *Nano Letters*, vol. 16, no. 7. pp. 4641–4647, 2016

[32] A. Fabas, Ouazzani EI, JP Hugonin, C Dupuis, R Haidar, JJ Greffet, P Bouchon, "Dispersion-based intertwined SEIRA and SPR effect detection of 2,4-dinitrotoluene using a plasmonic metasurface," *Opt. Express*, vol. 28, no. 26, p. 39595, 2020

[33] G. Lan, Z. Jin, J. Nong, P.Luo, C. Guo, Z. Sang, L. Dong, W Wei, "Narrowband perfect absorber based on dielectric-metal metasurface for surface-enhanced infrared sensing," *Appl. Sci.*, vol. 10, no. 7, 2295, 2020

[34] W. Liu, Z. Lai, H. Guo, and Y. Liu, "Guided-mode resonance filters with shallow grating," *Opt. Lett.*, vol. 35, no. 6, p. 865, 2010

[35] Z. Jacob, "Nanophotonics: Hyperbolic phonon-polaritons," *Nat. Mater.*, vol. 13, no. 12, pp. 1081–1083, 2014

[36] B. Zhao and Z. M. Zhang, "Perfect mid-infrared absorption by hybrid phonon-plasmon polaritons in hBN/metal-grating anisotropic structures," *Int. J. Heat Mass Transf.*, vol. 106, pp. 1025–1034, 2017

[37] N. K. Emani, T. F. Chung, A. V. Kildishev, V. M. Shalaev, Y. P. Chen, and A. Boltasseva, "Electrical modulation of fano resonance in plasmonic nanostructures using graphene," *Nano Lett.*, vol. 14,

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3190907

no. 1, pp. 78–82, 2014.
[38] W. Liu, B. Hu, Z Huang, H Guan, H Li, X Wang, Y Zhang, H Yin, X Xiong, J Liu, Y Wang., "Graphene-enabled electrically controlled terahertz meta-lens," *Photonics Res.*, vol. 6, no. 7, p. 703, 2018
[39] J. González-Colsa, G. Serrera, J. M. Saiz, F. González, F. Moreno, and P. Albella, "On the performance of a tunable grating-based high sensitivity unidirectional plasmonic sensor," *Opt. Express*, vol. 29, no. 9, p. 13733, 2021
[40] G. Lan, S. Zhang, H Zhang, Y. Zhu, Y. Qing, D. Li, J. Nong, W. Wang, L. Chen, W. Wei, "High-performance refractive index sensor based on guided-mode resonance in all-dielectric nano-silt array," *Phys. Lett. A.*, vol. 383, no. 13, pp. 1478–1482, 2019
[41] Y. L. Liao and Y. Zhao, "Ultra-narrowband dielectric metamaterial absorber for sensing based on cavity-coupled phase resonance," *Results Phys.*, vol. 17, no. March, p. 103072, 2020
[42] S.-F. Lin, CM Wang, TJ Ding, YL Tsai, TH Yang, WY Chen, JY Chang, "Sensitive metal layer assisted guided mode resonance biosensor with a spectrum inversed response and strong asymmetric resonance field distribution," *Opt. Express*, vol. 20, no. 13, p. 14584, 2012
[43] Q. Liu, G Cheng, Y Wang, "Fabrication of internal diffraction gratings in transparent materials using femtosecond laser," *Acta Photonica Sin*, vol. 33, no. 11, p. 1290, 2004.
[44] S. Sharma, R. Kumari, S. K. Varshney, and B. Lahiri, "Optical biosensing with electromagnetic nanostructures," *Rev. Phys.*, vol. 5, no. May, p. 100044, 2020
[45] S. Patel, J. Parmar, V. Sorathiya, R. Zakaria, T. K. Nguyen, and V. Dhasarathan, "Graphene-Based Plasmonic Absorber for Biosensing Applications Using Gold Split Ring Resonator Metasurfaces," *J. Light. Technol.*, vol. 39, no. 17, pp. 5617–5624, 2021
[46] L. Qian, K. Wang, W. Zhu, C. Han, and C. Yan, "Enhanced sensing ability in a single-layer guided-mode resonant optical biosensor with deep grating," *Opt. Commun.*, vol. 452, no. 101, pp. 273–280, 2019
[47] M. Jablan, H. Buljan, and M. Soljačić, "Plasmonics in graphene at infrared frequencies," *Phys. Rev. B - Condens. Matter Mater. Phys.*, vol. 80, no. 24, pp. 1–7, 2009
[48] M. Diem, PR Friffiths, JM Chalmers, "Vibrational spectroscopy for medical diagnostics"*Vibrational Spectroscopy for Medical Diagnosis Pharmaceutical Applications of Raman Spectroscopy, Wiley (2008)*.
[49] L. Kühner, M. Hentschel, U Zschieschang, H. Klauk, J. Vogt, C. Huck, H. Giessen, F. Neubrech,, "Nanoantenna-Enhanced Infrared Spectroscopic Chemical Imaging," *ACS Sensors*, vol. 2, no. 5, pp. 655–662, 2017
[50] J. Cheng, Y. Kuang, P. Gu, S. Feng, Y. Zhu, C. Tang, Y. Guo, Z. Liu, F. Gao, "Strong Magnetic Plasmon Resonance in a Simple Metasurface for High-Quality Sensing," *J. Light. Technol.*, vol. 39, no. 13, pp. 4525–4528, 2021
[51] B. Zhao and Z. M. Zhang, "Study of magnetic polaritons in deep gratings for thermal emission control," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 135, pp. 81–89, 2014
[52] J. Hu, W. Xie, J. Chen, L. Zhou, W. Liu, D. Li, Q. Zhan, "Strong hyperbolic-magnetic polaritons coupling in an hBN/Ag-grating heterostructure," *Opt. Express*, vol. 28, no. 15, p. 22095, 2020
[53] P. Li, M. Lewin, AV Kretinin, JD Caldwell, KS Novoselov, T Taniguchi, K Watanabe, F. Gaussmann, T Taubner, "Hyperbolic phonon-polaritons in boron nitride for near-field optical imaging and focusing," *Nat. Commun.*, vol. 6, no. May, pp. 1–9, 2015
[54] M. He, GR Iyer, S Aarav, SS Sunku, AJ Giles, TG Folland, N Sharac, X Sun, J Matson, S Liu, JH Edgar, "Ultrahigh-Resolution, Label-Free Hyperlens Imaging in the Mid-IR," *Nano Lett.*, vol. 21, no. 19, pp. 7921–7928, 2021
[55] A. N. Grigorenko, M. Polini, and K. S. Novoselov, "Graphene plasmonics," *Nat. Photonics*, vol. 6, no. 11, pp. 749–758, 2012,
[56] H. Song, S. Zhou, Y. Song, X. Wang, and S. Fu, "Tunable propagation of surface plasmon-phonon polaritons in graphene-hBN metamaterials," *Opt. Laser Technol.*, vol. 142, no. July 2020, p. 107232, 2021
[57] L. Ju, B. Geng, J. Horng, C. Girit, M. Martin, Z. Hao, HA Bechtel, X Liang, A. Zettl, YR Shen, F. Wang, "Graphene plasmonics for tunable terahertz metamaterials," *Nat. Nanotechnol.*, vol. 6, no. 10, pp. 630–634
[58] T. L. Myers, RG Tonkyn, TO Danby, MS Taubman, BE Bernacki, JC Birnbaum, SW Sharpe, TJ Johnson, "Accurate Measurement of the Optical Constants n and k for a Series of 57 Inorganic and Organic Liquids for Optical Modeling and Detection," *Appl. Spectrosc.*, vol. 72, no. 4, pp. 535–550, 2018
[59] H. Zhou, J. Zhu, Z. Liu, Z. Yan, X. Fan, J. Lin, G. Wang, Q. Yan, T. Yu, PM Ajayan, JM Tour, "High thermal conductivity of suspended few-layer hexagonal boron nitride sheets," *Nano Res.*, vol. 7, no. 8, pp. 1232–1240, 2014
[60] Y. Khalavka, C. Ohm, L. Sun, F. Banhart, and C. Sönnichsen, "Enhanced thermal stability of gold and silver nanorods by thin surface layers," *J. Phys. Chem. C*, vol. 111, no. 35, pp. 12886–12889, 2007
[61] R. C. Ropp, *Group 17 (H, F, Cl, Br, I) Alkaline Earth Compounds*, vol. 17. 2013.
[62] F. Ye, J. Lee, and P. X. L. Feng, "Electrothermally Tunable Graphene Resonators Operating at Very High Temperature up to 1200 K," *Nano Lett.*, vol. 18, no. 3, pp. 1678–1685, 2018
[63] X. Y. Fang, X. X. Yu, H. M. Zheng, H. B. Jin, L. Wang, and M. S. Cao, "Temperature- and thickness-dependent electrical conductivity of few-layer graphene and graphene nanosheets," *Phys. Lett. Sect. A Gen. At. Solid State Phys.*, vol. 379, no. 37, pp. 2245–2251, 2015

Rashmi Kumari received the M.Tech degree from IIT ISM Dhanbad, India in 2017. At present, she is working as a Research Scholar with the Department of E&ECE, Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India. Her research interests include two-dimensional materials, optical metasurafces, biosensing, switching, FEM and FDTD simulations.

Shubhanshi Sharma received M.tech degree from Delhi Technological University, Delhi, India in 2017. At present, she is working as a Research Scholar with the Department of E&ECE, Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India. Her research interests include metasurfaces, FEM and FDTD simulations.

Shailendra Kumar Varshney received the PhD degree from University of Delhi, Delhi, India in 2005. He is currently an Associate Professor with the Department of E&ECE, Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India. His research interests include PCF based optical devices, nonlinear dynamics in optical fiber and ring resonators, optical communication and metasurfaces.

Basudev Lahiri received the PhD degree from University of Glasgow, Scotland, UK in 2010. He is currently an Assistant Professor with the Department of E&ECE, Indian Institute of Technology (IIT) Kharagpur, Kharagpur, India. His research interests include metamaterials, nanofabrication, photonic crystals, early detection of Oral cancer.
.

© 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:31:43 UTC from IEEE Xplore. Restrictions apply.

Ecosystem Based Adaptation and Disaster Risk Reduction in Indian Himalayan Region of Darjeeling

Raju Sarkar[1], Saif Alam[2*], Sanchit Jain[3], and Sarthak Saharan[4]

[1]Department of Civil Engineering, Delhi Technological University
Bawana Road, Delhi, India

[2]Department of Civil Engineering, Delhi Technological University
Bawana Road, Delhi, India

[3]Department of Civil Engineering, Delhi Technological University
Bawana Road, Delhi, India

[4]Department of Civil Engineering, Delhi Technological University
Bawana Road, Delhi, India

Abstract. Mountain ecosystems are one of the major landforms in the world. Due to rampant industrialisation and urbanisation, mountainous ecosystem too has bore the brunt of environmental problems such as rising temperature and deglaciation. It has led to disruptions in hydrological functions, impact on forest and biodiversity and also human health implications. Ecosystem based adaptation is a nature based method which centres around environment related concerns and addresses the problem by reducing extreme weather conditions and thus enhancing sustainability and resilience. Darjeeling Himalaya is a mountainous region to the north-western side of the Indian state of West Bengal and has presence of numerous natural calamities such as Earthquakes, Landslides, Cloudbursts, etc. The local administration and the government has taken various measures and policies such as building bye laws, disaster management plan, draft development plan, limit building initiatives in order to act upon environment related hazards. With greater emphasis on EbA approaches to Eco-DRR in the regional policies of Darjeeling, mitigating the impact of climate can be drastically reduced which in turn will lead to Human and Economic Development. The EbA approach helps in (1) enhancing the community adaptive capacity or resilience, (2) improve or regenerate ecosystems to help increase value qualitatively and quantitatively of the produce for local communities and (3) integrating the current produce with market requirements,thereby increasing its economic viability.

Keywords: Ecosystem, Disaster Risk, Mountain Geography, Rural Areas, Climate Change

1. INTRODUCTION

The Himalaya and the contiguous sloping areas in the northeastern district of the domain of India is together called as Indian Himalaya Region (IHR). It addresses an exceptionally feeble and delicate Mountain Ecosystem [20]. The district, together called as Indian Himalaya Region (IHR), contains various ranges some of which surpass 7000 m in stature, high knolls, lake bowls, cold deserts, between montane valleys, profound crevasses, snowfields, ice sheets and alluvial fields [22]. A considerable lot of Asia's huge streams which includes Ganga, Indus and Brahmaputra and their endless feeders move through Himalayan reaches which give help to various human developments along their fertile valleys.

The Indian Himalayan region spans across many states which form the northern boundary of the Indian territory. The forest cover measurements made for the Himalayan states by the Forest Survey of India (FSI) shows that in numerous Northeastern hilly states forest cover has decreased in area during 1997 – 2001, although there is a slight recovery in following years [21].

State	Geographical Area	1997	1999	2001	2003	2005
Arunachal Pradesh	83,743	68,602	68,847	68,045	67,692	67,777
Himachal Pradesh	55,673	12,521	13,082	14,360	14,359	14,369
J&K	222,236	20,440	20,441	21,237	21,273	21,273
Sikkim	7,096	3,129	3,118	3,193	3,262	3,262
Uttaranchal	53,483	23,243	23,260	23,938	24,460	24,442

Fig.1 Distribution of Forest Cover in various states of IHR (in km^2)

Mountain districts have become inclined to catastrophe with a lopsidedly high occurrence of disasters in contrast with other kinds of environmental system [23]. These mountain ranges are geo-genuinely unsteady and dynamic construction and is exceptionally delicate to slant precariousness. Likewise, the district is set apart by a tough geology, delicate stone conditions, high seismicity, and furthermore high precipitation. These qualities have made the Himalayan ranges profoundly inclined to enormous scope development of structural plates, avalanches, and to the cycles of expulsion of surface. The Himalayas have likewise been exposed to high measure of human mediations. Exercises for work and advancement have developed and expanded the danger and weakness of the locale across the Himalayas.

2. LITERATURE REVIEW

The Himalayas give indispensable biological system administrations to the natural and monetary security of individuals living downstream[25]. They are a vault of geographical and agrarian resources and collected wild merchandise, other than being a focal point of social, social, and strict character to countless individuals (Saxena et al. 2001; Birch et al. 2014)[24]. Notwithstanding, the Himalayas are likewise one of the most corruption inclined and delicate mountain ranges on the planet. They are more defenseless because of land reasons, the pressure brought about by the expanding human populace, abuse of regular assets, and impacts of environmental change, which have diverse repercussions on human prosperity (Liu and Chen 2000; Dyurgerov and Meier 2005). The temperature increment is anticipated to be more perceptible in the Himalayan locale (INCCA 2010; Singh et al. 2011) with resultant effects like frosty melts, silt activation, dry spell, avalanches, and blaze floods (Dale et al. 2001; Wulf et al. 2012). These effects have brought about decreased usefulness of regular and human changed environments, loss of the travel industry income, framework and food and water security for the networks living in the area (WHO 2012; Badola et al. 2014; Smith 2014)[1].

The idea of environment administrations has turned into a significant model for connecting the working of biological systems with human prosperity (Fisher et al. 2009)[2]. In any case, even as interest for environment administrations is developing, human activities are reducing the ability of biological systems to fulfill these needs. Sound normal asset the board methodologies can frequently invert biological system debasement and upgrade the commitments of environments to human prosperity (MEA 2005). As of late, a huge and quickly developing assortment of examination

has been trying to recognize, portray and esteem environment labor and products (Bagstad et al. 2013) and plan ways of guaranteeing their constant arrangement[4]. Current biological system based methodologies have proposed that strategies and measures be embraced that consider the job of environment administrations in lessening cultural and natural weakness through multi-sectoral and staggered approaches (Andrade et al. 2011). Market analysts, tree huggers, and lobbyists have proposed different sorts of strategy blend, contending that 'solitary instrument' or 'single-procedure' approaches are off track and when utilized independently neglect to address complex reasons for environment administration debasement (for example Gunningham and Sinclair 1999; Ring and Schroter-Schlaak 2011)[3].

3. EbA and Eco-DRR

EbA centers around environment related regular perils, long haul means changes in environment and future vulner-abilities. For instance, it can prompt ascent in ocean levels and can add to change precipitation designs. These all includes an expansive arrangement of strategies to adjust to the environmental change. It includes biological system the board and the executives of its administrations to assist with lessening the weakness of individuals to the negative effects of the environmental change. For example, utilizing the technique for woodland insurance to assist with water maintenance in areas that are becoming drier. Eco-DRR (Disaster Risk Reduction) - It is a procedure for diminishing peril to typical risks by thwarting settlement and improvement in a disaster slanted districts by using an overall coordinated organic framework as a trademark pad to get people similarly as real properties in any area. Perils can join both climate as well as non-climate related dangers. For instance, tremors, volcanoes, landslides, waves. Inter-section of EbA and Eco-DRR-This incorporates the executives of environment hazard, implying environment related normal risks like tempests, floods, dry spell, avalanches, fire and so on for example mangrove reclamation or salt bogs rebuilding for waterfront assurance from floods of tempest.

3.1 Link between Disaster, Climate and Development

India's unique geo-climatic conditions leaves it exposed to a variety of disasters, including floods, cyclones, earthquakes, and drought. Around 58.6% of the country's landmass is sensitive to earthquakes; 12% of the landmass is subject to floods; and 68 percent of the country's agricultural area is vulnerable to droughts. A multitude of socio-economic issues, such as poverty, inequality, expanding populations, precarious livelihoods, and the growing economy's developmental priorities, all contribute to India's vulnerability to disaster risk. Direct losses from catastrophes in India cost alomost 2% of the India’s GDP and over 12% of the union government's income on average (World Bank 2003). Floods are responsible for annual losses in the United States, costing $7 billion per year (UNISDR 2015)[6]. Between 1980 and 2017, India saw 278 flood disasters, resulting in a total economic loss of $58.7 billion and affecting over 750 million people throughout the country (UNISDR & CRED 2018). The losses experienced by disasters, such as deaths, injuries, and destruction of buildings/infrastructure, are primarily included in these estimates of human and economic costs of catastrophes. The indirect effects of disasters on communities are frequently understated or unaccounted for. Hazard informed improvement orchestrating can pass on climate and disaster change shortcomings. There are different public ventures and plans which can go probably as potential area centers for mainstreaming fiasco and climate related concerns at both public and sub-public levels. These advancement activities can further develop strength of organizations against calamities and climatic risks by guaranteeing permission to economical work and fundamental administrations like clinical consideration, preparing, sanitization, water supply, etc; thus, diminishing weaknesses and updating limits of networks to withstand the fiasco and ecological change impacts[5].

4. MOUNTAIN ECOSYSTEM

It is broadly acknowledged that the debasement of mountain environments isn't any territorial issue yet a worldwide one. The Himalayas which are endowed with an extraordinary and quiet normal brilliance, plentiful assets, enduring waterways, other than being a storage facility of uncommon types of verdure, comprises one of the most undermined environments of the world. Given its delicate land profile, Himalayas is defenseless against numerous ecological perils and calamities of various kind, nature, and extent. Some among them incorporate quakes, avalanches, snow torrential slides, floods, trash streams, pandemics, and flames. These have been the most widely recognized reasons for death, harm, and obliterations in the Himalayan district [7-8].

4.1 Impact of Climate Change on Mountain Eco-System

The mountain ecological systems impact the air course altogether; however, they additionally show a lot of variety in the nearby climatic examples. In this manner, mountain biological system is probably going to get impacted the most in case there is any worldwide environmental change [10]. The effect of environmental change on mountain ecological system but can't be anticipated with sensible precision because of absence of long haul and timetable examinations [28]. In these areas even a slight change can possibly have a huge effect, especially in minimal conditions that are under pressure. A portion of the ordinarily noticed effects are:

4.1.1 Impact on Glaciers

Himalayan frosty masses have a locale around 23,000 km^2. It is most likely the greatest space of ice sheet set aside new water isolated from the polar (north and south) locale [11]. Chilly masses are an eventual outcome of climate and natural change. It is truly figured as certifiable thermometer of a risky air deviation. Audits coordinated for satellite imaging and ground assessments done by Indian Space Research Organization's Space Application Center (SAC) has set up the ice sheets have diminished from 2,077 km^2 to 1,628 km^2 in Himachal Pradesh alone – an overall deglaciation of 21 % in forty years. As demonstrated by SAC as much as 127 ice sheets of under 1 km^2 size have lost 38% of their geological locale beginning around 1962[27]. The greater ones, which are little by little getting separated into extra pieces, has withdrawn by however much 12% which is upsetting. Experiences on cold decline in IHR is open for only latest 150 years [13]. It reveals that the Himalayan frosty masses are pulling out at a typical speed of 18-20 m yr^{-1}. The pace at which some of the significant Himalayan ice sheets are retreating is introduced beneath:

Name of Glacier	Period(in years)	Recession (in m)	Average rate(m/yr)
Milam Glacier	108	1350	12.50
Pindari Glacier	121	2840	23.4
Gangotri Glacier	29	580	20.00
Tipra Bank Glacier	26	325	12.50
Dokriani Glacier	29	480	16.50
Chorabari	41	238	5.80
Shankulpa	76	518	6.80
Poting	51	262	5.13
Glacier no-3 Arwa	24	198	8.25
Bara Shigri	07	219	31.28
Chhota Shigri	03	54	18.5
Sonapani	52	899	17.2
Kolai	49	800	16.3
Zemu	07	193	27.5

Fig.2 Rate of Glacial Recession

4.1.2 Disruptions in Hydrological Functions

Expansion in encompassing temperature can have solid impact on neighborhood climate design. It is anticipated that there would be a vertical change in different climatic zones with slight expansion in temperature. Likewise, there is an expanding proof that colder time of year precipitation as snow fall has declined throughout the long term. Expansion in glacial mass softening is probably going to build overflow and frigid lake explosion floods. Decrease in the Cryosphere can likewise modify upstream hydrology, stream, essential efficiency, and mountain cultivating. Different outcomes of diminished hydrological capacities incorporate deficiency of drinking water, decrease in rural and hydropower creation. Effect of diminished catchment capacities in IHR is probably going to be undeniably more genuine on urban communities, cities and towns which lie on the downstream section including Himalayan lower regions that are dependent totally (or somewhat) on streams and waterways of these mountains [12].

4.1.3 Implication on Human Health

A portion of the consequences for human wellbeing are expanded encompassing temperature. It will then, at that point, cause warm pressure, bringing about inconvenience, physiological pressure and medical affliction. The recent concern of nature of water is presumably going to be furthermore weakened by ecological shift. The chances of spread of water-borne contaminations is going to augment. Without a doubt, even as of now, induction to safe drinking water is confined in the Himalayan Region. Ecological shift will in like manner impact in trans-mission of overwhelming ailments by dreadful little creatures, for instance vector-borne ailments like dengue, wilderness fever, and schistosomiasis. Such contaminations are vulnerable to the temperature similarly as to the land use variations [14].

4.2 Measures to Deal with Impacts

India is a signatory to the Kyoto Protocol. Thus it is resolved to abide by significant standards as referenced in the National Environmental Policy, NEP 2006[20]. However conformity to standards wherever should emphatically affect the mountain climate, the mountain individuals themselves should get very much adjusted and reasonably furnished to manage the resultant outcomes of environmental change[17]. This would require a progression of mechanical intercessions upheld by logical exploration and financial variations. It will be basic for accommodating the discernments on environmental change issues as seen by the specialists, strategy producers the mountain individuals, a larger part of whom are too poor to even consider making any such strides[16].

4.2.1 Establishing Networks of Meteorological Station Across the Region

Establish framework for environmental change research, particularly an organization in the Indian Himalayan Region of meteorological stations, and start incorporated review on the climatology through facilitated exertion among different foundations, gather specialized (determining, observing, planning ,and preparing for experts) political and monetary help for the said program[30]. Models should be created utilizing composite information to foresee the progressions and measure their effect in the different environments of the IHR for better administration procedures[15].

4.2.2 Increase Monitoring of Glacial Recession

Glacial masses in chosen areas inside the IHR, for example, those having fast deglaciation as of late should be checked on ceaseless reason for deciding pattern of the effect of worldwide environmental change. Stretch out examinations on ice sheets even to Eastern Himalaya[31].

4.2.3 Disease Surveillance

With the movements in the climate of mountain regions, a couple of vector-borne sicknesses will undoubtedly move to the higher heights. As such there is a sincere need to begin facilitated disease perception and deciding system for spread of vector borne sicknesses across the IHR[29].

4.2.4 Integration for Development of Mountain Region

Mountain is a significant wellspring of water, energy and organic variety. They are a wellspring of various assets like minerals, backwoods items and horticultural items. Likewise they are utilized for diversion reason. Mountain conditions are fundamental for the endurance of the worldwide biological system as they address mind boggling and interrelated environment of our planet. Mountain biological systems are, in any case, quickly evolving. They are defenseless against expanding soil disintegration, avalanches and quick loss of territory and hereditary variety. On the human side, there is unavoidable poverty among mountain tenants and besides there is loss of local data. Thusly, most overall mountain districts are experiencing environmental debasement. Thusly, genuine organization of mountain resources and monetary improvement of people justifies a brief movement[18-19].

State	2004-05	2009-10	2011-12
Arunachal Pradesh	31.4	25.9	9.20
Assam	34.4	37.9	31.98
Himachal Pradesh	22.9	9.5	8.06
Jammu & Kashmir	13.1	9.4	10.35
Manipur	37.9	47.1	36.89
Meghalaya	16.1	17.1	11.87
Mizoram	15.4	21.1	20.46
Nagaland	8.8	20.9	18.88
Sikkim	30.9	13.1	8.19
Tripura	40.0	17.4	14.05
Uttarakhand	32.7	18.0	11.26
India	37.2	29.8	21.92

fig.3 Percentage of Population below Poverty Line (Tendulkar Method): Indian Himalayan States

4.2.5 Addressing Poverty and Improving condition of women and children in Mountain Regions

Consideration of the fundamental reasons behind poverty in mountainous areas and women drudgery is of utmost urgency. It cannot be attended by looking everything through a conservationist glass. A balance must be struck between conservation of environment and development. Population which is below the poverty line in the mountain regions are greater than other parts of the country and its greatest impact is upon women and children in mountain areas, as majority of the able-bodied adults migrate out of these regions in search of livelihood. Also, the agriculture of these regions is at subsistence levels [26].

5. DARJEELING

The Himalayas in the Darjeeling is a rapidly urbanizing area in northeastern locale of India, continuously introduced to ordinary risks like shakes, torrential slides, and changing instances of precipitation in view of ecological change. This paper explores the marvelous underpinnings of fiasco danger creation in the region according to the viewpoint of cataclysm value. We the design in the Darjeeling Himalayas using the improvement of multi praised considerable

constructions as our object of assessment and considering data assembled from gatherings, plans and technique records, and part insight. This consistent assessment shows that a standard course of action of significant worth is a critical beginning stage for evaluating improvement processes and their commitment to catastrophe hazard, yet moreover illuminates how contemplations of calamity esteem are fascinating to express spots. By using such a construction, changed to fit explicit settings and conditions, we acknowledge that metropolitan improvement specialists can develop a more clear and taught technique for surveying the value with respect to their work.

5.1 Hazards and Climate Change

With its unsafe mountain inclinations and fast streaming streams, the Darjeeling Himalayas is introduced to different geologic and hydrologic hazards. During the tempest season, weighty precipitation produces streak flooding and triggers torrential slides, rockfalls, and other soil chances. Torrential slides, a trademark component of mountain conditions exacerbated by anthropogenic activities, consistently cost lives, upset system structures, and lead to enormous property mischief and land hardship. In 1968, for instance, just about 1000 millimeters of storm fell on the area in just three days, which made cataclysmic flooding and north of 20,000 torrential slides that killed enormous number of people and demolished homes, associations, and establishment like roads and platforms.

The district is also at high risk of tremors. As shown by the Geologic Survey of India, the Darjeeling Himalayas is by and large arranged in seismic zone IV, which has a bunch of encounters and partiality for gigantic and hurting events. In 2011, a 6.9 significance seismic quake with a point of convergence in bordering Sikkim killed 111 people and hurt or destroyed a couple thousand designs. Seismologists acknowledge that the district is sensible past due for a critical shudder of 8.0 significance or more noticeable, with conceivably disastrous results.

5.2 Risk

For the Darjeeling Himalayas, and for other urbanizing areas facing colossal risks from normal dangers, a significant request regarding catastrophes and progression has to do with the value of these risks. Many records of environmental value use the aftermath of unequivocal disasters, with the upside of knowing the past and the variations of the risk scene recognized, as their fundamental wellspring of verification. In this article, we battle that review the peril that stores up in the manufactured environment before disasters occur, to understand the value of the metropolitan progression processes by which that danger is made and shared.

5.3 Current Guidelines

District Nodal Officer (Environment): Division of Biodiversity and Environment has undertaking to make approaches, techniques, regulation, and rules to shield uprightness biodiversity, biological systems, and economical improvement in Darjeeling area [31].

The key jobs include Facilitation in creating arrangements, procedures, regulation, rules, plans and program in the region. Research and improvement. Capacity building and bringing issues to light on biodiversity, biological system, biosecurity, climate. Coordinate with concerned service and office for the administration of biodiversity, environments, and bio-wellbeing. Develop and fortify collaboration with improvement accomplices, common society, private area, the scholarly community, and other applicable partners to help preservation and the executives of biodiversity, biological systems, and biosecurity [25].

District Level Biodiversity Management Committee (DBMC) - Region level BMC is an approach making body reconstituted in January 2021 to empower feasible turn of events and furthermore to guarantee financial, natural, social, and social equilibrium inside Darjeeling region. DBMC involves 11 individuals alongside various squares, regions, and panchayats level Biodiversity Management Committees inside the locale. It is led by the District Magistrate, Darjeeling.

6. DISASTER MANAGEMENT

2030 been set as the end year for long-term activities in order to align the NationalDisaster Management Plan with post-2015 international accords. To synchronize the National Disaster Management Plan with the post-2015 peaceful accords, 2030 has been imagined as the closure year for long haul activities.The following agreements and steps are taken up.The government adopted three key international disaster management agreements in 2015:Sendai Framework for Disaster Risk Reduction (SFDRR) in 2015;Sustainable Development Goals (SDGs) (2015to30) in 2015;Paris Agreement on Climate Change at the UNFCCC 21st Conference of Parties (COP-21). Following documents are used NDMA and respective inputs from various state govts, SDMA;National Disaster Management Plan (NDMP) in 2016 and 2019;Dısaster Management Act(DMA) ,2005;Darjeeling disaster management plan has been extensively used.

The Indian Himalayan Region (IHR) stretches across 2,500 kilometres between the Indus and Brahmaputra River systems, and covers an area of about 5.3 lakh square kilometres.While glacial lake risks and distributions are studied in numerous glaciated places across the world, the Indian Himalayas have received comparatively little attention.The Himalayan region of India has a unique geographical entry into our nation, with a vast range of terrain, geology, soil, climate, flora and wildlife, and distinct ethnic groups with diverse socio-cultural traditions.

Risk mitigation techniques in the Himalayan area are primarily meant to mitigate or restrict the negative consequences of previously existing catastrophe risks using a mix of structural and non-structural measures.Once the area has been classified as hazard prone, it is critical that the government and the community take the steps outlined above. The mitigation strategy may differ depending on the key dangers identified, based on this methodology.

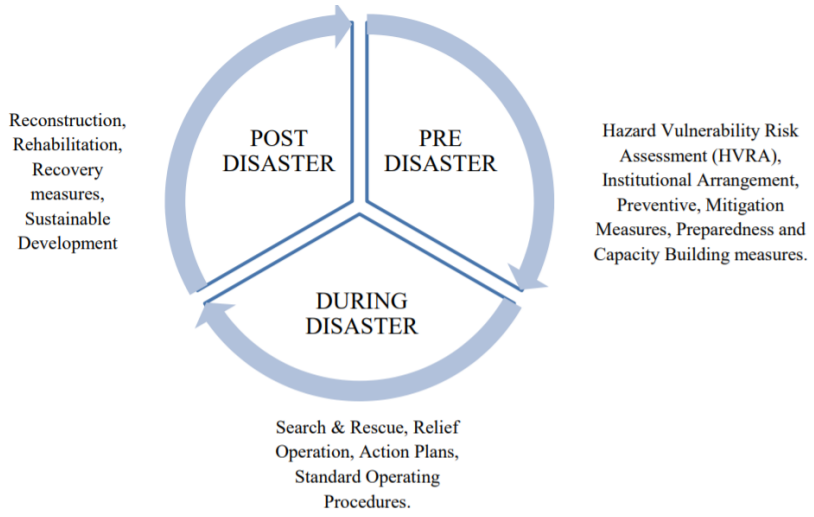


fig.4 : District Disaster management plan Darjeeling 2020-21

6.1.1 EARTHQUAKE

The Himalaya Mountains due to tectonic plate movement that has led to an increase in its height. Measures taken up ın Darjeelıng (wrt.earthquakes) as per the NDMA Act:District administration assured that multi-hazard-resistant measures are included in the development and implementation of social housing projects Ensuring that all applicable construction codes are followed.Strengthening and earthquake retrofitting as directed by the Ministries.Keeping a strict Vigil and safety audit of essential infrastructure and lifeline buildings in the district Ensuring effective implementation along with the incorporation of social audits in the district.

6.1.2 CLOUDBURST

Cloudburst is a short-term severe precipitation event that happens over a limited region and produces 10 cm or more rainfall in an hour.Raindrops are shipped upwards by the air current as opposed to falling as precipitation when immersed mists can't create downpour because of the vertical development of warm flows of air coming about to unnecessary buildup. Raindrops become too heavy to be transported aloft after a certain point and fall together in a flash.

As per NDMA, SDMA guidelines with respect to Cloudbursts, construction operations along river banks must be regulated, with specific attention paid to water levels after severe rains; Constricting and regulating water flow through strengthening embankments, barrages, and dams;Providing appropriate resources to re-energize the SDMA (funds and personnel);Localized planning that takes into account the region's ecological fragility and involves local populations;Preserving the purity of eco-sensitive zones by regulating infrastructure developments;IMD's improved forecasting and use of new technologies to monitor and anticipate extreme weather occurrences can help with early warning, evacuation, and preparedness;Disaster management and prevention are incorporated into the development planning process.ISRO's Space Application Centre (SAC) has created a model for heavy rainfall/cloud burst alarm to mitigate the effects of a cloud burst.It is now in pilot mode, but if its efficacy is established, it will go a long way toward saving lives.

6.1.3 LANDSLIDE

A landslide is a disaster phenomenon that occurs when a pile of rock, rubble, or dirt slides down a hill. NLRM (National Landslide Risk Management Strategy,2019) includes hazard mapping, monitoring, and early warning systems, as well as other areas of landslide catastrophe risk reduction and management.It involves landslide stabilisation and mitigation, as well as awareness campaigns, capacity building, training, legislation, and policies.Promotion and the adoption of landslide rehabilitation and mitigation strategies that are effective.Creating institutional capacity limit and preparing for geoscientists, specialists, and organizers is needed for viable administration of threat[31].

Followıng NDMA AND SDMA guıdelınes have been ıncorporated ın Darjeelıng disaster management plans such as:-Recharging the ground water and benefitting the environment by storing excess water in catchment regions to decrease the damage from flash floods.In the catchments, creation of dig runoff collection ponds. Plant fuel and fodder trees on all common lands. Planting grass cover in barren places, particularly on slopes, is an important component of an integrated watershed management programme[31]. Grazing should be strictly controlled. Better grasses can be planted when the region has been totally protected from grazing.Use surface vegetative cover to shield the ground from the pounding action of raindrops, bind soil particles, and slow the flow of water. Other darjeelıng specıfıc measures beıng taken up by the dıstrıct admınıstratıon such as the department undertakes the following functions:Upgradation and Preparation of the annual.DDMP;Directing the concerned departments and stakeholders on calamity preparedness measures; Releasing grants for the dead or ex gratia amount to the victims[32].

7. POLICIES INCORPORATING EbA and Eco-DRR

Disaster Risk Reduction outlines a huge piece of the current peril decrease measures [21]. The Yokohama Strategy (1994), the Indian Disaster Management Act (2005), Hyogo Framework for Action (HFA) (2005 2015) and Administrative Reforms Commission Report (2006) of the Government of India have ceaselessly settled the significance of calamity decline. In accordance with these frameworks and rules, the Darjeeling area has also accepted a movement of measures to diminish risk in the spaces under its ward. The activities can immediately be assigned follows:

Building Byelaws that gives a comprehensive managerial framework to the advancement of constructions inside Darjeeling common district. Disaster Management Plan targets propelling a "culture of availability" through proactive disaster the chiefs' extents of "orchestrating, status and expectation" for making "fiasco flexible Darjeeling with

outright risk decline as the essential noticing limit in each developmental theory and drives to ensure viable new development [27]."

Draft Development Plan is a significant instrument for consolidated arrangement and changed development and gives a functional technique towards organizing the disaster the board stresses with progression orchestrating. Limit Building Initiatives Capacity building is a huge piece of cataclysm preparation. Concerning the Darjeeling Municipal district, the Municipality, NGOs, Civil Defense Department, and various social events have taken measures to improve and build the neighborhood with respect to better fiasco status. These social occasions, in any case, don't work in isolation; rather they have hoped to enhance their undertakings through cooperation and linkages. The Municipality, in a joint effort with the Civil Defense Department and NGOs having a few aptitudes in fiasco the chiefs like Red Cross and Anugyalaya and samaj, has begun measures for limit collecting that fuse honing the neighborhood the calamity, planning for rescue and help and developing neighborhood through courses and studios [25].

8. CONCLUSION

Eco-DRR and EbA play an important role, so as to reduce the vulnerability to climate change of any particular region. The frameworks discussed above thus proposes an approach for the facilitation of evaluation and monitoring of interventions to help fight various inequalities, risks associated with climate change. Thus, in this hour, it becomes even more important to study Disasters, risks associated with them along with their mitigation. It indicates that a comprehensive methodology needs to be adopted for design, implementation, monitoring, evaluation, and review of such interventions. The frameworks discussed above are not expected to be prescriptive in nature and therefore can be applied flexibly as per the requirement and conditions in any given region.

Based on the analysis above of the policies and procedures in place in Darjeeling, it does not fully accommodate the upcoming environmental hazards and Darjeeling should make a standardized department for EbA and Eco-DRR related enquiries. Pilot studies for execution of both the concepts should take place region wise and collective data should be used in making new action plans for betterment of the ecosystem of Darjeeling. With the implementation of Ecosystem based approaches, it will benefit various sectors. Due to such policies in the regions of Darjeeling, water, agriculture, forests sectors will be conserved which will further lead to improvement in energy, health and tourism. Due to these advantages of EbA, it is imperative to explore these opportunities and invest for further study on ecosystem-based approaches to adaptation. It demands inputs of cross sectoral policy makers for policy formulation to achieve an increased level of sustainability and resilience in the mountain ecosystem.

References

[1] S. Besky, “The Land in Gorkhaland,” *Environ. Humanit.*, vol. 9, no. 1, pp. 18–39, 2017, doi: 10.1215/22011919-3829118.
[2] B. Zutshi, *Manipur flood 2015: Preparedness and disaster risk reduction strategies*. 2018.
[3] S. Singh, S. M. Tanvir Hassan, M. Hassan, and N. Bharti, “Urbanisation and water insecurity in the Hindu Kush Himalaya: Insights from Bangladesh, India, Nepal and Pakistan,” *Water Policy*, vol. 22, pp. 9–32, 2020, doi: 10.2166/wp.2019.215.
[4] R. Seidler and K. S. Bawa, “Ancient risks, current challenges in the Himalayas,” *Econ. Polit. Wkly.*, vol. 51, no. 41, pp. 63–67, 2016.
[5] S. Bera, B. Guru, R. Chatterjee, and R. Shaw, “Geographic variation of resilience to landslide hazard: A household-based comparative studies in Kalimpong hilly region, India,” *Int. J. Disaster Risk Reduct.*, vol. 46, no. December 2019, p. 101456, 2020, doi: 10.1016/j.ijdrr.2019.101456.
[6] E. Brink *et al.*, “Cascades of green: A review of ecosystem-based adaptation in urban areas,” *Glob. Environ. Chang.*, vol. 36, pp. 111–123, 2016, doi: 10.1016/j.gloenvcha.2015.11.003.
[7] T. K. Datta, “Advantages , Constraints And Key Success Factors In Establishing Origin- And Tradition-Linked Quality Signs : The Case Of Darjeeling Tea , India,” *Evolution (N. Y).*, 2009.
[8] A. CHHETRI and L. TAMANG, “Decentralization of Water Resource Management: Issues and Perspectives Involving Private and Community Initiatives in Darjeeling Town, West Bengal,” *Ann. Natl.*

Assoc. Geogr. India, vol. 39, no. 2, pp. 240–255, 2019, doi: 10.32381/atnagi.2019.39.02.6.

[9] L. Dorren and M. Schwarz, *Ecosystem-Based Disaster Risk Reduction and Adaptation in Practice*, vol. 42, no. December. 2016.

[10] F. C. Godoi, S. Prakash, and B. R. Bhandari, "Final report Final report," *Rev. 3D Print. potential red meat Appl.*, no. 23 February 2021, pp. 1–61, 2021.

[11] S. Bhattacharya, G. Ghosh, T. Banerjee, S. Goswami, and P. Das, "Socio-Environmental Survey of an Ecologically Important Hamlet of Darjeeling District West Bengal, India," *Int. Lett. Nat. Sci.*, vol. 33, pp. 51–72, 2015, doi: 10.18052/www.scipress.com/ilns.33.51.

[12] B. K. Shah, V. Kedia, and R. K. Jha, "Integrated Vendor-Managed Time Efficient Application to Production of Inventory Systems," *Proc. 6th Int. Conf. Inven. Comput. Technol. ICICT 2021*, pp. 275–280, 2021, doi: 10.1109/ICICT50816.2021.9358504.

[13] P. Chaudhary *et al.*, "Consistency of local perceptions of climate change in the Kangchenjunga Himalaya landscape," *Curr. Sci.*, vol. 101, no. 4, pp. 504–513, 2011.

[14] J. A. Klein *et al.*, "An integrated community and ecosystem-based approach to disaster risk reduction in mountain systems," *Environ. Sci. Policy*, vol. 94, no. December 2018, pp. 143–152, 2019, doi: 10.1016/j.envsci.2018.12.034.

[15] H. Sandhu and S. Sandhu, "Linking ecosystem services with the constituents of human well-being for poverty alleviation in eastern Himalayas," *Ecol. Econ.*, vol. 107, pp. 65–75, 2014, doi: 10.1016/j.ecolecon.2014.08.005.

[16] M. Ghosh and S. Ghosal, "Living with insecurity: A critical examination of the environmental problems linked with social problems in rural west bengal, india," *Rom. J. Geogr.*, vol. 63, no. 1, pp. 53–64, 2019.

[17] A. Shroff, B. K. Shah, A. Jha, A. K. Jaiswal, P. Sapra, and M. Kumar, "Multiplex Regulation System with Personalised Recommendation Using ML," *Proc. 5th Int. Conf. Trends Electron. Informatics, ICOEI 2021*, pp. 1567–1573, 2021, doi: 10.1109/ICOEI51242.2021.9453005.

[18] B. K. Shah, V. Kedia, R. Raut, S. Ansari, and A. Shroff, "Evaluation and Comparative Study of Edge Detection Techniques," vol. 22, no. 5, pp. 6–15, 2020, doi: 10.9790/0661-2205030615.

[19] B. K. shah, A. Kumar and A. Kumar, "Dog Breed Classifier for Facial Recognition using Convolutional Neural Networks," *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 2020, pp. 508-513, doi: 10.1109/ICISS49785.2020.9315871.

[20] J. P. Basu, "Climate change vulnerability and perception of different occupational groups of household in the hill regions of Darjeeling district, West Bengal," *Indian J. Ecol.*, vol. 47, no. 1, pp. 248–258, 2020.

[21] J. A. Rather and M. S. Bhat, "….. A …. Research Periodic Evaluation of Long Term Precipitation Trends in Eastern Middle Himalayas : A," no. December, 2020.

[22] B. P. Resurrección *et al.*, *The Hindu Kush Himalaya Assessment*. Springer International Publishing, 2019.

[23] S. Besky, "Rural Vulnerability and Tea Plantation Migration in Eastern Nepal and Darjeeling," *Himal. Res. Pap. Arch.*, pp. 1–21, 2007, [Online]. Available: http://hdl.handle.net/1928/3304.

[24] R. Seidler, "Seidler et al . 2018 - Progress on integrating climate change adaptation and disaster risk reduction in South Asi ...," 2018.

[25] B. K. Shah, A. K. Jaiswal, A. Shroff, A. K. Dixit, O. N. Kushwaha, and N. K. Shah, "Sentiments Detection for Amazon Product Review," *2021 Int. Conf. Comput. Commun. Informatics, ICCCI 2021*, 2021, doi: 10.1109/ICCCI50826.2021.9402414.

[26] E. Sharma *et al.*, *The Hindu Kush Himalaya Assessment*. Springer International Publishing, 2019.

[27] Reliefweb.int, "Using mobile phone technologies for Disaster Risk Management: Reflections from SHEAR, June 2021 - World | ReliefWeb," no. June, pp. 1–29, 2021, [Online]. Available: https://reliefweb.int/report/world/using-mobile-phone-technologies-disaster-risk-management-reflections-shear-june-2021.

[28] M. Kumar, A. Bhatia, P. Gupta, K. Jha, R. K. Jha, and B. K. Shah, "S IGN L ANGUAGE A LPHABET R ECOGNITION U SING," no. Icices, pp. 1859–1865, 2021.

[29] A. Snyder, "Shortage in the Mountains of Plenty : Water Supply in Mountain and Hill Cities throughout the Hindu-Kush Himalayan Region," 2014.

[30] M. Chamling and B. Bera, "Spatio-temporal Patterns of Land Use/Land Cover Change in the Bhutan–Bengal Foothill Region Between 1987 and 2019: Study Towards Geospatial Applications and Policy Making," *Earth Syst. Environ.*, vol. 4, no. 1, pp. 117–130, 2020, doi: 10.1007/s41748-020-00150-0.

[31] V. Lo, *Synthesis report on experiences with ecosystem-based approaches to climate change adaptation and disaster risk reduction*, vol. Technical, no. 85. 2016.

[32] V. Kedia, S. R. Regmi, K. Jha, A. Bhatia, S. Dugar, and B. K. Shah, "Time Efficient IOS Application for CardioVascular Disease Prediction Using Machine Learning," *Proc. - 5th Int. Conf. Comput. Methodol. Commun. ICCMC 2021*, no. Iccmc, pp. 869–874, 2021, doi: 10.1109/ICCMC51019.2021.9418453.

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

Effects of Different Laser Shape Pulses on Laser Wakefield Acceleration

Vikrant K Patel[1], Raunaq Singh[2], Nipul Agarwal[3], Monika Yadav[4], Suresh C Sharma[5]

[1,2,3,4,5] *Department of Applied Physics, Delhi Technological University, New Delhi, India*
vikrantkp172000@gmail.com, raunaqsingh10@gmail.com, nipulagarwal09@gmail.com, raomonika77@gmail.com, suresh321sharma@gmail.com

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835736

***Abstract*— The authors of this work discuss the impact of varying laser shape pulses on laser Wakefield acceleration. Researchers have identified a relativistic plasma wave with an extraordinarily high electric field, which has the potential to be used to accelerate plasma particles in general, as well as for particle acceleration in general, according to the researchers. The acceleration process is fueled by a massive accelerating field gradient created by plasma wave, which acts as the primary motivator for acceleration process to take place and accelerate the plasma wave. A plasma wave provides the energy for the acceleration process. To ignite this kind of plasma wave, it is possible to produce an ionizing wakefield with a single laser, and the plasma wave will then interact with the laser itself. A wiggler magnetic field is utilized in this article to illustrate that augmentation of electron acceleration created by plasma wave may be investigated by deploying a plasma wave generator, which is described in detail elsewhere. A faster forward motion is required to make use of the higher resonance provided by the wiggler field, thus, electrons caught in the plasma wave are propelled forward at a faster pace. As a consequence, the amount of energy that the electrons accumulate throughout their acceleration increases dramatically.**

Keywords—Laser; Laser Wakefield; Acceleration

I. INTRODUCTION

It has been some years since significant advances in both the quality and the energy of the particle beams that have been made accessible have been seen. The use of radio frequency (RF) waves has unquestionably been the most dependable and successful method of particle acceleration up to this point in history.

Accelerating electrons to relativistic energy may be costly; the next generation of particle accelerators will likely be several kilometres long and cost billions of dollars. However, using the massive (hundreds of GigaVolts per metre) electric fields in laser-produced plasmas to accelerate electrons might be one approach to lower the size and expense of future accelerators for use in both basic physics research and technology applications. These fields are hundreds of times stronger than traditional accelerators' accelerating fields. [9]

When it comes to particle acceleration, even while the employment of radio frequency waves is still the most effective approach for some applications, this is no longer true for the purest uses of accelerators due to the increasingly high energy barrier that is being encountered in this sector. A shaped laser pulse, which is a kind of LWFA, is suggested in this work as a means of investigating. Since the refractive index varies with different materials, the laser pulse's form changes and become skewed.

It is possible to detect the acceleration gradient in the laser-plasma accelerator is about a thousand times greater than the acceleration gradient in the standard accelerator when they are compared side by side. Due to their high energy, electrons may move at speeds greater than the speed of light over short distances (on the scale of a few centimeters and with energies more than a few hundred Megavolts (MEV).

The refractive index lowers, resulting in an increase in phase velocity. In order to compensate for this reduced leading edge phase velocity, a pulse compression happens prior the laser pulse [10], causing the pulse compression to take place prior to the laser pulse [10]. When high-intensity pulses are used to generate plasma waves, the waves get nonlinearly steepened as a result. Plasma waves condense into a cavity-like structure as they approach the wave's center, and the change in refractive index inside this structure is very modest. Hence, the majority of laser energy is pushed forward in front of the cavity, where it is redshifted and slides fast back into the wave frame, increasing the pulse steepening at its front and lowering it at its rear as it does so.

Keeping high-power laser technology growing and improving at a rapid pace is critical to ensuring the continued improvement of LWFA performance in the years to come. As a result of the great efforts made over the preceding few years to enhance the peak power of the laser, which had previously been unexplored. LWFA examinations have been conducted throughout the course of the last ten years utilising a range of different kinds of PW lasers, with the most recent tests carried out in 2012. The most recent experiments were carried out in 2012, with the most recent experiments carried out in 2012.

The accelerating particles in a diverse range of applications [1, 2], has been one of the most notable unintended consequences of particle acceleration in synchrotrons. Molecular crystallography and fluorescence research are just two of the topics covered in this course, which also include topics such as chemical analysis and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

medical imaging. Despite the fact that synchrotrons need just a few GeV of particle energy to generate X-rays, they are still only suitable for use in large national laboratories, so limiting their use and, as a result, their accessibility. A new way of X-rays for scientific investigation, the free electron laser (FEL), is now accessible and may be used in conjunction with other devices to enhance the scientific investigation. It engaged with the electromagnetic wave in such a way that the resultant radiation increases across the wave's propagation length, a phenomenon that is characterized by an increase in the total amount of radiation emitted.

The strong laser pulse radiation pressure causes plasma electrons to be shifted ahead and to the side, resulting in the formation of positively charged ion wakes, which are fields that are behind and traveling in parallel with the laser pulse radiation pressure. If you look at Figure 1, you will see that once laser pulse has been passed, the displaced electrons are drawn back and trapped by positively charged wakefield, which causes them to accelerate longitudinally as a consequence of the positive charge. In order to characterize this unique kind of particle acceleration, the term "laser wake eld acceleration" (also known as LWFA) was created to describe it).

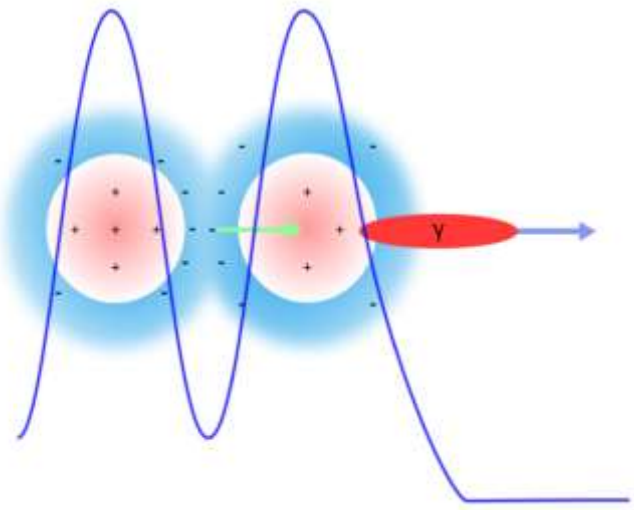


Figure 1. LWFA [1]

(Figure 1). As a result of this, attraction and acceleration, electrons in the wake are drawn to the positively charged wakefield and accelerate as a result of this attraction and acceleration. We can determine the amount of ionization present in the plasma by drawing a blue line over it. Energy obtained over time is shown in Figure 2, which shows the cumulative quantity of energy gained over time.

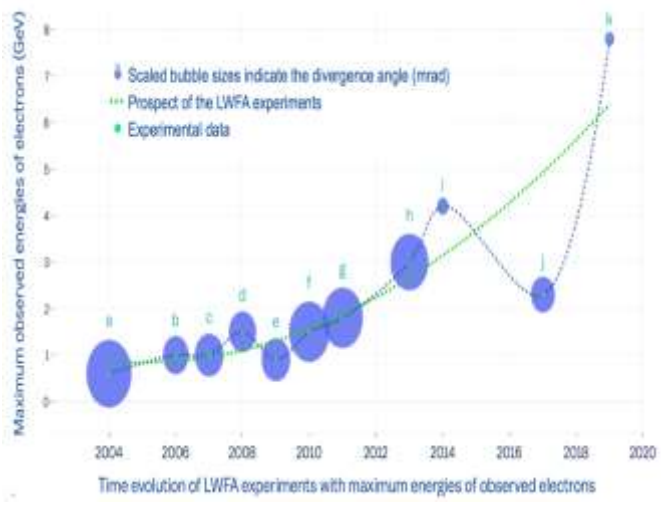


Figure 2. Maximum energies versus time evolution [3]

When a low intensity, brief laser pulse interacts with a main target material, most of the laser energy is absorbed by the target's free electrons owing to inverse Bremsstrahlung. The difference between the lattice temperatures and electron (Te > Ti) in a transitory nano-equilibrium state may explain this outcome. The electrons are cooled despite a minor energy transition between the lattice and electron heat conduction. Time t is substantially larger than e for picosecond laser ablation (t >> te). [34]

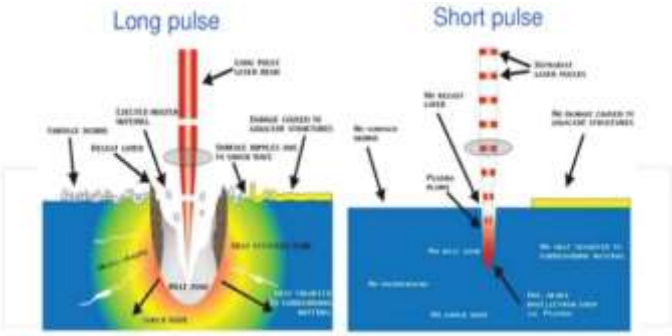


Figure 3. Long pulse laser and short pulse interaction with material [34]

II. HIGH QUALITY BEAM

This part provides a whole discussion of the fundamental principle of (LWFA). As a surfer takes up energy from ocean wave, electrons in laser produced plasma may takes energy from the wave in plasma. When a strong laser pulse travels through the plasma and displaces background electrons due to laser's ponderomotive force (or light pressure), these waves are created. Behind the laser pulse, these waves travel at almost light speed. Electrons in the background plasma may be trapped and accelerated to extremely high energy over very short distance by the longitudinal electric fields of plasma waves with sufficiently great amplitude.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

Using a laser pulse that is both very rapid and strong, it is possible to build coherent wakefields with relativistic amplitude and high phase velocity. This is possible as long as the plasma provides stable support for the laser pulse during its travel through plasma. The wakes and sheaths of the plasma are compared and contrasted, utilizing comparison and contrast to compare and contrast the plasma architecture. High-amplitude wakefields create collective resonant oscillations of electron eigenmode over the whole plasma, but because of their rapid wake phase velocity c and the superfast nature of laser pulses, wake stability and stiffness are induced in the plasma electrons. When the phase velocity falls, sheaths are generated, and when the phase velocity increases, wakefields are formed, and the same is true in the other direction. And it is these dynamics that we are interested in understanding. The results of a slew of experiments undertaken all around the world suggest that this theory is rapidly becoming a reality, both in terms of the promise of high-energy accelerators and the likelihood of larger applications in the not-too-distant future.
Experimental evidence has shown that a major increase in the quality of electron beam generated may be achieved by experimenting with a large number of various injection processes. The bubble/blow-out regime, density gradient injection, colliding laser pulses injection, ionization injection, and recently, longitudinal injection are some of the technologies that are used in this industry.
It is possible to boost the energy of particle accelerators via the use of laser wakefield acceleration and plasma wakefield acceleration; however, putting these concepts into practice is very difficult to do so successfully. The "laser-plasma wakefield accelerator," which was developed by an international team of researchers, brings together the advantages of both methods of accelerating electrons. If such accelerators are built, it is possible that the number of high-energy particle physics facilities throughout the globe will expand. Alternatively, they may be used to create free-electron lasers, which will be beneficial in a wide variety of scientific fields in the years to come.

Despite the fact that particle accelerators have been responsible for some of the most significant scientific discoveries, the need to smash particles with increasing levels of energy is testing the limits of current technology. Even while the use of electric fields in classical accelerators is ubiquitous, the danger of component damage by arc discharge, or even simply melting them as a result of overheating, limits the greatest field that may be used. A substantial number of physicists who work in the area are unable to access accelerator facilities as a effect of the increasing size and cost of these facilities. To be sure, the globe is home to a small number of cutting-edge particle-physics research facilities – such as LHC(Large Hadron Collider) at CERN – that are capable of conducting cutting edge particle physics research at the greatest possible levels of complexity and efficiency.

However, even though particle accelerators have been responsible for some of the most important scientific discoveries, the demand to smash particles with ever-increasing intensity is placing an increasing strain on the capabilities of the equipment that is already on the market today. Electric fields are used in traditional accelerators, but the danger of arc discharge or even simple melting of components due to overheating prevents the application of a large enough field. A consequence of this has been a significant increase in the cost and size of accelerator facilities, making them inaccessible to the vast majority of physicists and other scientific researchers. Sure, there are several world-class particle physics research facilities, such as the Large Hadron Collider at CERN, that are capable of doing ground-breaking particle physics research.

Wakefield accelerators, which are akin to the wake left by a ship plowing through water, generate an electric-field wave in a stationary plasma by delivering a pulse of energy to it. This wave, which is comparable to the wake produced by a ship plowing through water, may be utilized to accelerate particles in a similar manner. If a trailing cluster of particles is timed correctly, they may be able to surfer this wave, resulting in a far greater acceleration than would otherwise be feasible in a conventional accelerator system. On the contrary, it is the development of this burst of energy that poses the most difficult task.

In light of recent developments in particle physics research, the plasma wakefield accelerator concept seems to be one of the most promising possibilities for the construction of the next-generation particle-physics collider. Before this can be achieved, however, there are a number of hurdles that must be cleared out of the way.

It is important to first design an acceleration mechanism for positrons, which is currently being researched. This is due to the fact that the PWFA is more effective for electrons than it is for positrons, and as a result, the acceleration system for positrons must be built first. According to current knowledge, this is the case because of the decelerating force imposed by the plasma electrons on the positrons, which is responsible for the acceleration. It is also necessary to acknowledge that the accelerated particles are acceptable for collision in order for them to collide in the first place in order for them to clash at all. There are significantly more stringent criteria for the TeV plasma wakefield accelerator, which makes it far more difficult to meet them. Plasma wakefield accelerators had an energy level in the tens of GeV range at the time of this writing, which was amazing. A few examples of the possible applications for this technology include material imaging, radiotherapy, and radiolysis in biological systems, to name a few.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

III. LITERATURE REVIEW

Because of their importance in the search for a more complete understanding of the fundamental concepts of particle physics and their applications, particle accelerators have grown in importance in interactions between very small fissures of invisible matter. Particle accelerators are also capable of testing the interactions and governing forces between very small fissures of invisible matter. Additionally, particle accelerators may examine the forces and interactions between extremely tiny fissures in the visible matter by examining the interactions and governing forces between the fissures when it comes to physical size. Its diameter is 27 kilometers, making it the largest particle accelerator in the world. Since the proton beams are being accelerated to practically the speed of light at the LHC, their energy has increased to around 7 TeV as a result. A large number of nations from all across Europe are collaborating to improve scientific research. Currently, under construction, the Future Circular Collider (FCC), which will be a larger accelerator and collider than the LHC and will be operational after the LHC's lifespan is through, will eventually replace the LHC. It was revealed in December 2018 that the Federal Communications Commission (FCC) would be housed in a tunnel measuring 100 kilometers in length [1], according to the Federal Communications Commission's conceptual design study. Electric field gradients greater than 100MVml have the potential to destroy the accelerator structure itself, which is one of the reasons larger accelerators are being constructed in this circumstance. It is evident that the size of current accelerators must be expanded in order to create more energy for charged particles; as a result, conventional accelerators are becoming larger and more expensive [2]. It is necessary to research novel acceleration processes in order to overcome the limitations of the size-energy relationship, which is currently unsolvable.

They are equivalent to those achieved by conventional radiofrequency accelerators. Traditional radio frequency accelerators, on the other hand, need long tunnels of many kilometers in length to achieve the same gradient [3].

Veksler pioneered the concept of collective acceleration via the use of plasma in 1956, which was subsequently adopted by NASA. The advent of plasma as an accelerating medium also provided him with the opportunity to relieve the accelerator physicist of the challenges connected with the electrical breakdown of media, which had previously resulted in saturation of the accelerating gradient's upper limit. In the course of the subsequent investigation, which was done in an effort to test his concept, it was revealed that his notion also caused new challenges. Given that nonrelativistic electrons are necessary in order to capture nonrelativistic ions, the electron-triggered plasma wave must also become nonrelativistic (i.e., it must become "reflexing"). As a result, Mako and Tajima8) developed their own model in order to have a good understanding of phenomenon's nature. Their findings revealed that the plasma wave breakdown results in ion energy gains that are just a few orders of magnitude more than the electron energy gains, which is in direct opposition to what Veksler anticipated by M/m.

For the purpose at producing high energy electron beams, laser wakefield acceleration (LWFA), a technique for creating accelerating field structures formed by a strong, ultrashort pulsed laser traveling through a dense plasma, is utilized. It has been demonstrated experimentally3,4,5,6,7,8,9 that plasma accelerators are capable of producing accelerating gradients that are 1000 times stronger than those ejected by conventional sources1,2, and that the development of high-energy electron beams is possible as a result of the absence of a breakdown limit. 11, 12, 13, 14, 15, 16, and 17). With the betatron X-ray beams generated by LWFA11, it has been shown that high-resolution tomographic imaging may be achieved. The fact that they have low divergence (of the order of few milliradians18) and short duration makes them particularly well-suited for use in a wide variety of applications in engineering, medical science, public safety, and homeland security, as well as research and technology11–17 is one of their most attractive characteristics (of the order of less than 100 fs19). As previously proven, a single laser shot may be utilized to create betatron X-ray beams with micrometer size and keV energy. This demonstrates the potential of betatron X-ray sources for imaging complex objects in real time using high repetition rate laser systems.

Tajima and Dawson10 were the first to suggest the notion of Laser Wakefield Acceleration, which was originally launched in 1979 and has since gained widespread acceptance. (LWFA). It is possible to produce plasma wave by injecting a brief, intense laser pulse with suitable pulse duration and density into a plasma with the proper density. When this method is performed, a plasma wave is created that is referred to as an eigenmode of the plasma itself. The fact that this resonant wave may have a significant amplitude (which we refer to as relativistic energy amplitude) has been shown for the first time, leading to the electron motion in the wakefield collecting relativistic energy, as previously stated. In spite of the fact that this wave travels at a relativistic phase velocity, it is a structurally very powerful and long-lasting wave with an extended duration. Due to the increased difficulty of electron overtaking (and hence the prevention of wave breaking) and the increased visibility of electron concentration, a nonlinear object known as a density cusp singularity is formed that is out of phase with the rest of the system. When dealing with nonrelativistic phase velocity, we attempted to produce the LWFA at relativistic phase velocity in order to minimize the fragility of the wave, and we were successful in this endeavor. We were able to overcome the restricted trapping breadth of ions that was caused by the heavy mass of the particle, as previously discussed8, by accelerating electrons that were employed to accelerate the particles themselves. Although this latter difficulty becomes critical when we talk about laser ion acceleration, we will not go into depth about it here; instead, we will refer you to a recent study paper on the subject that you may read about elsewhere for additional information on

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

the subject. 10. The Veksler plan's safeguards should be sufficient to deal with the issues that it faced throughout its implementation.

Because the accelerated focusing phase of the wakefield occupies just a portion of plasma wavelength, the injection strategy has a significant impact on the quality of electron beam that is created. As consequence of acceleration, it is also possible for the high-density accelerated electron bunch to generate its own wakefield as a result of its acceleration. A common word used to describe this phenomenon is the phrase "beam loading" (see also [2–4]). Due to fact that it allows electron bunch to become more compact, this process is critical in accelerating the growth of the energy spectrum of accelerated electron bunch. This is important in development of the energy spectrum because it allows the electron bunch to become more compact. Allowing the beam loading process to take its course on its own may result in a rapidly fluctuating field across the whole bunch length, which may result in the unwanted expansion of energy spreads throughout the full bunch length if allowed to continue.

When the injected bunch is properly formed, the electric field is flattened across the length of the bunch, resulting in a more even distribution of energy across the bunch as a consequence. As a result, a thorough knowledge of this phenomenon will be important in the development of the next generation of laser plasma accelerators, which will be capable of producing high quality electron beams in future.

In spite of fact that wakefield generated by the ponderomotive force of the laser pulse is independent of laser pulse polarization, it is seldom investigated if laser pulse polarisation has an influence on it. In fact, it is rare that it is investigated.

This feature, which takes advantage of shrinking dimensions and cost reductions, allows us to reduce the cost and size of future accelerators on a broad scale by decreasing their dimensions and costs. Taking into consideration the optimistic experimental findings obtained so far, it seems that the energy frontier may be stretched in the near future, with accelerators becoming more compact structures and experiments that can be performed on a tabletop becoming more routine. It is also important to emphasize that one of the major advantages of employing a plasma medium to accelerate tiny charged particle is that the structure of plasma state has already been reduced to a more manageable level, which is a significant benefit. This is due to the presence of plasma surrounding the accelerated particles, which prevents the particles from causing damage to the accelerator and necessitates their evacuation from the system. A recent study identified plasma-based acceleration systems (also known as plasma-based accelerator systems) as one of the most viable long-term replacements for the present generation of radiofrequency accelerators in the contemporary world.

Using a burst of energy to generate an electric-field wave in a stationary plasma – akin to the wake left by a ship plowing through water – wakefield accelerators are a revolutionary new technology that can be used to accelerate particles. Depending on how well a trailing cluster of particles timing their arrival to the wave, they may be able to surfer it, resulting in a far higher acceleration than would be feasible in a conventional accelerator. This, on the other hand, is a difficult job that requires much effort.

In recent findings, a highly promising acceleration regime known as the bubble regime [11] has been discovered, which results in the production of quasi-monoenergetic electron beam with a large amount of charge. The identification of this regime was accomplished by 3D Particle-In-Cell (PIC) simulations, which were carried out in 2002 and reported in 2002. [12] 3D PIC simulations of electron beam quality in the blow-out phase revealed a comparable gain in electron beam quality when the laser-intensity was reduced by factor of two, which was consistent with the findings of previous studies. Any of those regimes allows the laser energy to be concentrated in a region of the plasma whose radius is less than the radius of the laser while the laser is functioning. Whenever there is an application of the ponderomotive force, electrons in the plasma are forced to evacuate inward, leading to the formation of a positively charged cavity that is surrounded by an extremely high density of electrons just behind the laser beam. Because all of the injected electrons exhibit comparable starting characteristics when viewed in phase space, this suggests that the injection was successful. Due to the fact that the injection is tightly localized and that all of the injected electrons have comparable initial properties, this is the case. As previously observed [13-15], when the charge of the electron beam reaches a certain threshold, experimental tests revealed that the trapping process is terminated, resulting in the creation of an electron beam with quasi-monoenergetic distribution. Electrons are trapped far behind the laser during this phase [15], where they are no longer able to interact with the laser's electromagnetic field due to result of their imprisonment. As electron beam travels through the system, the quality of the beam improves as a result of the increased movement. To detect stable and regulated quasi monoenergetic electron beams for the first time, two counterpropagating laser beams [7] were used in conjunction with a collision system [8] in 2006. The results were published in 2006. The outcomes of the study were published in 2006 due to result of the research. When first laser pulse, also known as the injection pulse, collides with the second laser pulse, which is also known as the injection pulse, it is used to excite the wakefield, and it is used to heat electrons when second laser pulse, also known as the injection pulse, collides with the first laser pulse, also known as the injection pulse.

Putting the technologies of laser wakefield acceleration and plasma wakefield acceleration into practice, on the other hand, is a very difficult undertaking. A multinational team of researchers has now successfully accelerated electrons using both laser and plasma technologies, showing the advantages of both approaches. The researchers used a "laser-plasma wakefield accelerator" to achieve this feat. It is probable that due to result of this, the number of high-energy particle physics facilities throughout the globe may increase in the future. Free electron lasers, which are

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

employed in a wide range of scientific fields and may be manufactured in enormous numbers, can also be created with the help of such accelerators.

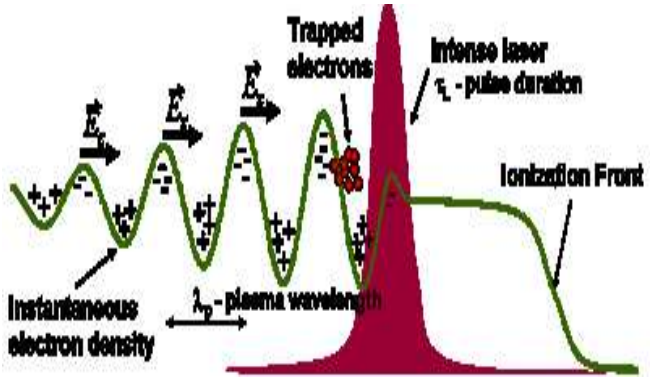


Fig. 4 Ionization [9]

As seen in Fig. 4, the ionization process is underway. This paper describes the invention of a scaling method for electron energy optimization, which will be useful for future research and development on electron accelerators.

IV. APPLICATIONS

It has been more than three decades since the concept of laser wakefield accelerators (LWFAs) was first discussed. While LWFAs has the potential to enable the construction of compact high energy particle accelerators, they will also provide scientific community with novel light sources that are currently unavailable. In an LWFA, a high intensity laser pulse is introduced into plasma, resulting in production of an electromagnetic wave as a result of laser pulse, which is subsequently returned to the source of laser pulse. This technology allows electrons to be collected and accelerated to energies in the GeV range, and it is now being used on a regular basis in a wide range of applications, including semiconductor manufacturing.

This article covers an analysis of light sources created from relativistic electrons emitted by LWFAs, which span from terahertz radiation to gamma rays in frequency, as well as their potential applications in a variety of disciplines. Assuming the existence of relativistic electrons, it is proposed that betatron motion, Compton scattering, and undulators are all approaches for producing X-rays or gamma-rays and that these techniques are all connected to one another. There is a common theoretical underpinning for all of their experimental findings, which is oscillating relativistic electrons in wakefield formed by a laser pulse, a counter-propagating laser field, or an undulator. Bremsstrahlung and terahertz radiation are two examples of light sources that are based on LWFA, and both are classed as light sources by the World Health Organization. By accelerating particles at very high speeds, plasma-based accelerators, also known as laser-wakefield accelerators (LWFAs), are able to produce incredibly dense ultra relativistic electron beams. Because of large acceleration gradients found in LWFAs, it is conceivable to produce ultra-relativistic electron beams using these devices. This distinction is given to LWFAs since they are devices that are used to accelerate electron beams to very high energies, which is why they are called LWFAs. A brilliant X-ray pulse of micrometer source size is generated when electrons traveling at high speeds pass through the intense focus fields of the wakefield. This pulse is visible in visible light because it is generated by accelerated electrons passing through the wakefield's intense focusing fields. Because of the intense focusing fields created by this pulse, it has the potential to be used for imaging applications. When processing heterogeneous materials, nondestructive X-ray phase-contrast imaging and tomography techniques may be utilized to determine the processing parameters of the materials as well as their structural composition and mechanical performance.

V. CONCLUSION

The following article covers a few laser wakefield acceleration-based laser effects that are presently being investigated and evaluated as a consequence of this. In some circumstances, such as when a wiggler magnetic field is used in its best form, it is feasible to get the highest amount of electron energy by using the cyclotron resonance. In order to find the optimal magnetic field strength and plasma wave amplitude for each circumstance under consideration, it is necessary to compute the energy gain of electrons in every situation. Using an algorithm that maximizes the amplitude of the plasma wave, the wiggler magnetic field, and beginning momentum of electrons while maintaining the initial position of the electrons constant throughout the experiment, as described previously, we were able to determine how much energy was gained by the electrons. With an increase in the amplitude of plasma wave, there is a big rise in amount of electron energy that may be produced. The electrons' velocity must be concentrated and stabilized in order for them to stay in the accelerating zone of the magnet, and this can only be done by concentrating and stabilizing the electrons' velocity. Due to consequence of acceleration generated by plasma wave, the electron gains more energy at a given point in time. But even while electrons with starting velocity close to speed of light in vacuum may interact resonantly with the plasma wave and wiggler magnetic field, the energy gain relative to the starting velocity is much greater when compared to the beginning velocity.

REFERENCES

[1] M. Yadav, S. C. Sharma, and D. N. Gupta, "Electron Acceleration through a Relativistic Electron Plasma Wave in an Inverse-Free-Electron Laser Mechanism," IEEE Transactions on Plasma Science, vol. 46, no. 7, July 2018, pp. 2521–2527, doi: 10.1109/TPS.2018.2843362.

[2] Ahmet Guzel & Altan Cakir (2019). A Literature Review on Plasma Wakefield Acceleration.

[3] K. Gopal, D. N. Gupta, and H. Suk, "The Effect of Pulse Length on Electron Laser Wakefield Acceleration by Skewed Laser Pulses," IEEE Transactions on Plasma Science, vol. 49, no. 3, March 2021, pp. 1152-1158, doi: 10.1109/TPS.2021.3059643.

[4] A. Abada, M. Abbrescia, S. S. Abdus-Salam, I. Abdyukhanov, J. Abelleira Fernandez, A. Abramov, M. Aburaia, A. O. Acar, P. R. Adzic, P. R. Adzic, P. R. Adzic, P. R. Adzic, P. R. Adzic, P. R. Adzic, P

[5] R. Bingham and R. Trines Plasma accelerator basics. Yellow Reports from CERN, 2016; 1(0):67.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

[6] Tajima, T., and Dawson, J. M. Laser-based electron accelerator. Phys. Rev. Lett., 43:267270, July 1979.

[7] E. Esarey, P. Sprangle, J. Krall, and A. Ting A concise summary of plasma-based accelerator concepts. IEEE Transactions on Plasma Science, 24(2):25228 (April 1996).

[8] E. Esarey, C. B. Schroeder, and W. P. Rev. Mod. Phys., 81:1229–1285, Aug 2009.

[9] S. P. D. Mangles, C. D. Murphy, Z. Najmudin, A. G. R. Thomas, J. L. Collier, A. E. Dangor, E. J. Divall, P. S. Foster, J. G. Gallacher, et al. Monoenergetic beams of relativistic electrons are generated as a result of strong laser plasma interactions. 535–538, 2004. 431(7008):535–538, 2004. 431(7008):535–538, 2004. Nature, 431(7008):535–538, 2004.

Hyung Taek Kim, Ki Hong Pae, Hyuk Jin Cha, I Jong Kim, Tae Jun Yu, Jae Hee Sung, Seong Ku Lee, Tae Moon Jeong, and Jongmin Lee. [10] Hyung Taek Kim, Ki Hong Pae, Hyuk Jin Cha, I Jong Kim, Tae Jun Yu, Jae Hee Sung, Seong Ku Lee, Tae Moon Jeong, and Jongmin Lee. Hyung Taek Kim, Ki Hong Pae, Hyuk Jin Cha, I Jong Kim, Tae Jun Yu, Jae Hee Sung, Seong Ku Lee, Tae Moon Jeong, and Jongmin Lee. [10] Hyung Taek Kim, Ki Hong Pae, Hyuk Jin Cha, I Jong Kim, Tae Jun Yu, Jae Hee Sung, Seong Ku Lee, Tae Moon Jeong, and Jongmin Lee. Electron energy enhancement, Physical Review Letters, vol. 111, no. 16, October 2013).

[11] Hyung Taek Kim, V. B. Pathak, Ki Hong Pae, A. Lifschitz, F. Sylla, Jung Hun Shin, C. Hojbota, Seong Ku Lee, et al. Stable multi-GeV electron accelerator driven by waveform-controlled PW laser pulses. Scientific Reports, August 2017, vol. 7, no. 1.

[12] Y. Glinec, S. Kiselev, S. Gordienko, E. Lefebvre, J. P. Rousseau, F. Burgy, and V. Malka. A laserplasma accelerator is used to generate a monoenergetic electron beam. 431(7008):541-544, 2004.

[13] Cs. Toth, W. P. Leemans, B. Nagler, A. J. Gonsalves, K. Nakamura, C. G. R. Geddes, E. Esarey, and C. B. Schroeder. GeV electron beams from a centimetre-scale accelerator. 2006, Nature Physics, 2(10), pp. 696-699.

[14] et al., S. Karsch, J. Osterho, A. Popp, T. P. Rowlands-Rees, Z. Major, M. Fuchs, B. Marx, R. Horlein, K. Schmid, and L. Veisz. Acceleration of electrons in a gas-filled capillary discharge waveguide on the order of GeV. The November 2007 edition of the New Journal of Physics contains the following article: 9(11):415-415.

[15] Nasr A. M. Hafz, Tae Moon Jeong, Il Woo Choi, Seong Ku Lee, Ki Hong Pae, Victor V. Kulagin, Jae Hee Sung, Tae Jun Yu, Kyung-Han Hong, and Tomonao Hosokai; Stable generation of gev-class electron beams from self-guided laserplasma channels. 2:571 EP, Nature Photonics, 2008.

[16] C. E. Froula, K. A. Marsh, C. P. J. Barty, L. Divol, R. A. Fonseca, S. H. Glenzer, C. Joshi, W. Lu, et al. Measurements of electron self-injection critical power in a laser wakefield accelerator. Physical Review Letters, vol. 103, November 2009 (21).

[17] Clayton, C. E., Ralph, J. E., Albert, F., Fonseca, R. A., Glenzer, S. H., Joshi, C., Lu, Marsh, K. A., Martins, S. F., and Mori, W. B. Self-guided laser wake eld acceleration surpassing 1 GeV is shown using ionization-induced injection. Physical Review Letters, September 2010 issue, 105 (10).

[18] Haiyang Lu, Mingwei Liu, Wentao Wang, Cheng Wang, Jiansheng Liu, Aihua Deng, Jiancai Xu, Changquan Xia, Wentao Li, and Hui Zhang; Acceleration of electron beams more than 1 GeV utilizing a laser wakefield from an ablative capillary discharge waveguide. 99(9):091502, August 2011. 99(9):091502, August 2011. Applied Physics Letters, 99(9):091502, August 2011.

[19] W.P. Leemans, A. J. Gonsalves, H.-S. Mao, K. Nakamura, C. Benedetti, C. B. Schroeder, Cs. Toth, J. Daniels, D. E. Mittelberger, and S. S. Bulanov Multi-GeV electron beams created in the self-trapping regime through capillary discharge-guided subpetawatt laser pulses Physical Review Letters, vol. 113, December 2014 (24).

[20] A. J. Gonsalves, K. Nakamura, J. Daniels, C. Benedetti, C. Pieronek, T. C. H. de Raadt, S. Steinke, J. H. Bin, et al. A petawatt laser directs and accelerates electron beams to 8 GeV in a laser-heated capillary discharge waveguide. Phys. Rev. Lett., 122:084801, February 2019.

[21] C. E. Clayton, D. Umstadter, C. Joshi, and C. Darrow. Collinear Optical mixing is used to generate relativistic plasma waves. Phys. Rev. Lett. 54:2343–2346, May 1985.

[22] Clayton, C. E., Marsh, K. A., A. Dyson, M. Everett, A. Lal, Leemans, W. P., Williams, R., and Joshi, C. Laser-excited relativistic electron plasma waves accelerate injected electrons at ultrahigh gradients. Jan 1993, Phys. Rev. Lett., 70:3740.

[23] N. A. Ebrahim. Lasers, plasmas, and particle accelerators: The twenty-first century's cutting-edge technology for particle acceleration. Phys. Canada, 45:178–188, 1989.

[24] Vargas, M., Schumaker, W., He, Z.-H., Zhao, Behm, K., Chvykov, B. Hou, Krushelnick, A. Maksimchuk, and Yanovsky, V. Stability, divergence, and energy spread of laser wakefield accelerated electron beams improved with the use of three-dimensional printed two-stage gas cell targets. Applied Physics Letters, 104, 2014 (17).

[25] S. Wang, C. E. Clayton, B. E. Blue, E. S. Dodd, K. A. Marsh, W. B. Mori, C. Joshi, S. Lee, P. Muggli, T. Katsouleas, F. J. Decker, M. J. Hogan, R. H. Iverson, P. Raimondi, D. Walz, R. Assmann X-ray emission caused by the wiggler betatron motion of a plasma. 2002, Phys. Rev. Lett., 88:135004.

[26] Antoine Rousse, Kim Ta Phuoc, Alexander Pukhov, Eric Lefebvre, Victor Malka, Sergey Kiselev, Frederic Burgy, Jean-Philippe Rousseau, and Donald Umstadter. In a relativistic laser-plasma interaction, synchrotron radiation is employed to create a keV x-ray beam. 2004; Phys. Rev. Lett. 93:135005.

[27] E. Esarey, B. A. Shadwick, P. Catravas, and W. P. Leemans Synchrotron radiation is produced when electron beams are focussed via plasma-focused channels. 65:056505, 2002, Phys. Rev. E.

[28] I. Kostyukov, A. Pukhov, and S. Kiselev A phenomenological theory for laser-plasma interaction in the bubble domain. Plasma Physics, vol. 11, no. 11, 2004.

[29] T. Katsouleas, M. Zhou, M. Tzoufras, F. S. Tsung, W. B. Mori. T. Katsouleas, W. Lu, C. Huang, M. Zhou, M. Tzoufras, F. S. Tsung. A nonlinear theory of multidimensional relativistic plasma waves' wakefields). Plasma Physics, vol. 13, no. 5, 2006

[30] C. Aniculaesei; V.B. Pathak; H.T. Kim; K.H. Oh; B.J. Yoo; E. Brunetti; Y.H. Jang; C.I. Hojbota; J.H. Shin; et al.

[31] Utilization of up-ramp plasma density profiles in a laser wakefield accelerator to increase electron energy. Sci. Rep., vol. 9, no. 1, pp. 1–7, 2019. [CrossRef]

[32] V. Tomkus; V. Girdauskas; J. Dudutis; P. Gecys; V. Stankevic; G. Raciukaitis; I. Gallardo González; D. Guénot; J.B. Svensson; Electron beams and betatron radiation are accelerated by a laser wakefield from multijet gas targets. [34] L.T. Ke; K. Feng; W.T. Wang; Z.Y. Qin; C.H. Yu; Y. Wu; Y. Chen; R. Qi; Z.J. Zhang; Y. Xu; et al. Electron Beams at Near-GeV Density-Tailored Plasma from a Laser Wakefield Accelerator at a Few Per-Mille Level 2021, 126, 214801 Phys. Rev. Lett.

[33] Guidance Optical in Meter-Scale PlasmaWaveguides B. Miao; L. Feder; J. E. Shrock; A. Goffin; and H. M. Milchberg 2020, Phys. Rev. Lett., 125, 074801.

[34] Abubaker Hassan Hamad, "Effects of Different Laser Pulse Regimes on the Ablation of Materials for Production of Nanoparticles in Liquid Solution", september 2016, DOI: 10.5772/63892

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:11:19 UTC from IEEE Xplore. Restrictions apply.

Environment, Development and Sustainability
https://doi.org/10.1007/s10668-022-02596-w

Energy input–output analysis and greenhouse gas emission in okra and tomato production in Chotanagpur plateau region of India

B. Sarkar[1] · B. Das[1] · P. K. Sundaram[1] · S. S. Mali[1] · A. P. Anurag[1] · A. Upadhyaya[1] · N. Chandra[1] · B. P. Bhatt[2] · A. Kumar[3]

Received: 29 October 2020 / Accepted: 25 July 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract
The information on greenhouse gas (GHG) emission and energy use patterns from vegetable production in the Chotanagpur plateau region of Jharkhand state is minimal. The current study analyzed the energy input–output and GHG emission and their relationship with the productivity of two critical vegetables grown in the region, viz. summer season okra and tomato. In this regard, data were collected from 30 vegetable farmers of the region in a pretested questionnaire through personal interviews. The results show that the overall energy input used in the okra production was 8828.71 MJ ha^{-1} of which human energy (27.62%), petrol (27.31%), farmyard manure (13.59%), and animal energy (13.22%) contributed the primary inputs. The total energy required for tomato production was 4798.66 MJ ha^{-1}, where petrol (25.13%) contributed the highest, followed by fertilizer (16.94%), diesel (14.67%), electricity (12.06%), farmyard manure (12.03%), and human energy (11.65%), respectively. The energy ratio (energy output to energy input) for okra and tomato was estimated at 2.85 and 7.58, respectively. The benefit cost ratios for tomato and okra production were 7.87 and 1.71, respectively, showing that the cultivation of both the vegetables is remunerative in the region, with tomato being more remunerative than okra. The total GHG emission was 875.41 and 322.75 kg CO_2eq ha^{-1} for okra and tomato, respectively. The economical use of inputs could help reduce GHG emissions in vegetable production.

Keywords Energy analysis · Okra · Tomato · Greenhouse gas emissions · BC ratio · Chotanagpur plateau

1 Introduction

The Chhotanagpur Plateau in Eastern India spans over most of Jharkhand state as well as parts of states, such as Odisha, West Bengal, Bihar, and Chhattisgarh. The region is dominated by tribal, mainly engaged in agriculture and wage activities (Census, 2011). Over

✉ B. Sarkar
bikas_sarkar2003@yahoo.com

Extended author information available on the last page of the article

70% of Jharkhand's rural population relies on agriculture and allied industries for their livelihoods. In this region, out of total cultivable area of 37.78 lakh ha, the net sown area of the state is only 25.57 lakh ha with cropping intensity at 116% and the main crop being paddy covering 80% of total cropped area (NABARD, 2017).

Favorable agroclimatic conditions in the Chhotanagpur plateau allow the cultivation of high-value varieties throughout the year. The two main summer vegetable crops grown in the area are okra and tomato. Okra (Abelmoschus esculentus L. *Moench*) is an annual vegetable crop cultivated in its tropical and subtropical regions. The agroecological condition of Jharkhand state favors successful cultivation of okra during spring, summer, and rainy seasons. To increase the production per unit area with the available resources, okra intercropping can successfully be adopted during the rainy season under various agroclimatic conditions of the plateau region of Jharkhand (Kandeyang et al., 2010); Doab region of Punjab (Saini et al., 2020); Chiplima in Odisha (Patra et al., 2021).

Similarly, tomato (*Solanum lycopersicum* L.), a native of tropical America, is considered the second most important remunerative solanaceous vegetable after potato. After China, India is the second largest tomato producer, with a global share of 18.9% and 11.2% in region and volume, respectively (Vanitha et al., 2013a). India has 882.0 thousand hectares of land under tomato cultivation, with 18,735.9 thousand metric tons and a productivity of 21.2 t ha^{-1} (Saxena & Gandhi, 2015). Jharkhand has a remarkable share in tomato production, having 22.3 thousand hectares of land under this crop with a production of about 401.6 thousand metric tons (Vanitha et al., 2013b). In Jharkhand, tomato is extensively cultivated in Ranchi, Lohardaga, Hazaribagh, and Godda districts covering approximately 13.9% of the area under vegetable cultivation. Given the better economic returns over paddy, the area under vegetable production has shown increasing trends over past few years (NABARD, 2017).

Agriculture in Jharkhand is transitioning from the traditional low energy input farming methods to a higher level of energy input methods of agricultural production to meet the food requirements of ever-increasing population. Today, agricultural production is dependent on the consumption of fossil fuels (Taskin and Vardar, 2016). Efficient use of energy will lead to the same level of development through less energy use and emissions (Ozturk et al., 2018). Tomato and other vegetables under foothill conditions of Eastern Himalayas help in improving productivity, profitability, and energy use efficiency (Kumar et al., 2019).

In agriculture, energy usage per unit area is a direct product of inputs like diesel, electricity, machines, seeds, fertilizers, and chemicals (Hamedani et al., 2011). The supply of agricultural inputs consumes much electricity and emits GHG leading to adverse environmental impacts (Kastner et al., 2012). Many agricultural operations such as field preparation, planting, fertilization, spraying, and refining are carried out using fossil fuels. CO_2 from burning of fossil fuels in agricultural equipment is emitted into the atmosphere (Lal, 2004; 2020). Extreme utilization of fertilizers is significant contributor to non-point pollution from agricultural supplies (Liu et al., 2013), environmental pollution (Ju et al., 2007), reduction in the value of cultivated harvests (Huang et al., 2006), and upsurge of harmful air emissions (Wang et al., 2008). With recurrent attack by invasive insect species and consequent crop damages, application of insecticides is growing at faster rate (Pimentel & Pimentel, 2008). The amplified usage of insecticides contributes to significant environmental pollution throughout the world (Toan et al., 2013).

Various indices like energy production, specific energy efficiency, energy productivity, and energy viability can be used to analyze the energy gains and environmental impacts of crop production systems (Khoshroo et al., 2021). Increased energy production would

encourage sustainable agronomy by minimizing environmental issues and avoiding the depletion of natural resources (Ghorbani et al., 2011).

Moreover, adoption of green energy sources and improved energy conservation will substantially contribute to sustainable energy production (Streimikiene et al., 2007). Therefore, considering the above facts, there are no available reports indicating energy consumption and estimation of GHG emissions of summer vegetable production in the Chotanagpur plateau region. The study's objectives focus on estimating energy use and evaluating GHG emissions for summer vegetable production.

2 Materials and methods

2.1 Experimental site

The sites selected for the present study were Malti and Kharsidag villages of Namkum block in the Jharkhand state of India. The selected villages are located at an average elevation of 629 m above the mean sea level, having hot and dry subhumid climatic with hot and dry summers and cold winters. The study villages are bounded between the latitude of 23° 15′ 07″ to 23° 16′ 14″ and longitude of 85° 22′ 41″ to 85° 24′ 32″. The study villages receive an average annual rainfall of 1295 mm, 82.8% received during June–September. The annual evaporation is about 1960 mm. The region has cold and humid winter months, with some inadequate rainfall during September and October, typically bright, sunny days, and cold nights. The soil in the villages was deep, well-drained, sandy loam with 65.9% and 21.7% silt, 12.4% clay, and 0.41% organic carbon with an acidic pH of 5.17 in the top 20 cm soil depth. The selected sites' average soil water content varies from 18.8 to 21.2% and 9.5 to 11.6% at field capacity and wilting point values. The topsoil's bulk density (0-20 cm) ranged between 1.42 and 1.65 g cm^{-3}.

2.2 Data collection

The data were collected through a direct consultation with farmers through a pretested schedule. The schedule included all kinds of inputs used to calculate agricultural activities, i.e., human, animal, fuel, electrical, seed, FYM, fertilizers, chemicals, machinery, and irrigation. During the study, energy input–output for the production of two summer vegetables, i.e., okra and tomato, was taken into account. Summer okra (Syngenta okra hybrid OH 597 and Arka Nikita) seeds were sown in the month of the 2nd week of June at a seed rate of 18 kg ha^{-1} with a spacing of 45 (row-to-row) × 20 (plant-to-plant) cm. The harvesting of okra started after 45–50 days of sowing. For summer tomato (variety Swarna Sampada) production, one-month-old healthy seedlings were transplanted at a spacing of 60 × 45 cm in the first week of July. The 1st harvesting was done after 60 to 80 days of transplanting of seedlings.

2.3 Estimation of energy balance

The energy budgeting of each vegetable crop was calculated using input energy in fertilizers, seeds, plant protection chemicals, fuels, human labor, animal power, machinery, and other agro-inputs. Direct and indirect energy inputs have calculated the energy inputs of

each crop. Direct energy inputs include humans, livestock, diesel, electricity, and irrigation. Indirect energy sources comprise energy used in the processing of crops and other agro-inputs, such as seeds, fertilizers, FYM, chemicals, and machinery. Energy sources have been listed as sustainable and nonrenewable. Renewable energy included human energy, animal energy, crop, FYM, and irrigation, while nonrenewable energy sources include diesel, gas, pesticides, fertilizers, and equipment. The production energy was provided in terms of grain and leaves, vines, and straw from vegetable yields. By multiplying the conversion coefficients for various devices and energy input sources, inputs have been converted from physical units to energy units (Tables 1 and 2). The energy indicators like energy use efficiency, energy productivity, and specific energy were calculated based on equations reported by various researchers (Canakci & Akinci, 2006; Devasenapathy et al., 2009; Esengun et al., 2007; Hamedani et al., 2011; Nassiri & Singh, 2009; Nourani & Bencheikh, 2019; Pishgar-Komleh et al., 2012; Tuti et al., 2012).

2.4 Energy balance

The analysis of energy input–output flows in the production process of okra and tomato was carried out to compare these distinct systems. Energy input flow was directly linked to input intensiveness of the system. Energy flux was computed considering crop management

Table 1 Energy equivalents of inputs and outputs in agricultural production. *Source*: Devasenapathy et al. (2009), Mittal et al. (1985), Taki et al. (2013) Tuti et al. (2012)

Particular	Unit	Equivalent energy (MJ unit^{-1})
Human labor		
Adult men	Hour	1.96
Women	Hour	1.57
Bullock pair (medium)	Hour	10.10
Diesel	Litre	56.31
Petrol	Litre	48.23
Electricity	kWh	11.93
Water for irrigation	m^3	1.02
N	kg	60.60
P_2O_5	kg	11.10
K_2O	kg	6.70
Farm yard manure (FYM),(dry mass)	kg	0.3
Plant protection		
Inferior chemicals (Cow urine and neem extract)	kg	10.00
Seed		
Okra	kg	25.60
Tomato	kg	1.0
Output		
Okra	kg	1.9
Tomato	kg	0.80
Leaves, vines and straw from vegetables (dry mass)	kg	10.0

Table 2 Energy coefficient (MJ h^{-1}) of various farm equipments. *Source*: Nassiri and Singh (2009)

Power source	Tool/equipment	Energy coefficient (MJ h^{-1})
Manual	Khurpa	0.031
	Spade	0.314
	Sprayer	0.502
Animal	Plough	0.627
	Harrow and planker	1.881
Tractor	Tractor (35 hp)	10.944
	Cultivator	3.135
	Electric motor	0.216
	2-wheeler (motor cycle)	2.73 (calculated)

and quantity of biomass production. Computations included both operational (direct) and nonoperational (indirect) sources of inputs. Operational energy inputs include manual work, fuel, and farm machinery, whereas nonoperational energy involves seed, fertilizer, and plant protection chemical.

2.4.1 Input energy

Total energy input to the system was calculated as the sum of energy from all the components of inputs used in the crop production process.

$$\mathrm{IE} = \sum_{i=1}^{n} \left(C_i\right) \tag{1}$$

where IE = input energy, C_i = energy of ith component (MJ ha^{-1}).

2.4.2 Output energy

The biomass crop yield provides the main product (grain) and by-product (straw/leaves/stalk/vine). Grain/seed and straw/haulm energy output was determined by multiplying the respective energy equivalent.

$$\text{Net energy return (MJ ha}^{-1}) = \text{Energy output} - \text{Input energy} \tag{2}$$

$$\text{Energy ratio} = \frac{\text{Total output energy } \left(\text{MJ ha}^{-1}\right)}{\text{Total input energy } \left(\text{MJ ha}^{-1}\right)} \tag{3}$$

$$\text{Energy productivity (kg MJ}^{-1}) = \frac{\text{Crop yield } \left(\text{kg ha}^{-1}\right)}{\text{Input energy } \left(\text{MJ ha}^{-1}\right)} \tag{4}$$

$$\text{Specific energy } \left(\text{MJ kg}^{-1}\right) = \frac{\text{Total input energy } \left(\text{MJ ha}^{-1}\right)}{\text{Yield of main crop } \left(\text{kg ha}^{-1}\right)} \tag{5}$$

$$\text{Energy profitability (PE)} = \frac{\text{Net enegy return } \left(\text{MJ ha}^{-1}\right)}{\text{Input energy } \left(\text{MJ ha}^{-1}\right)} \quad (6)$$

$$\text{Human energy profitability (HEP)} = \frac{\text{Output energy } \left(\text{MJ ha}^{-1}\right)}{\text{Labour energy } \left(\text{MJ ha}^{-1}\right)} \quad (7)$$

2.5 Greenhouse gas emission (GHG) estimation

The technique for calculating greenhouse gas (GHG) emissions in various farm activities is comparable to the energy study. In the energy analysis, contributions are translated into MJ and GHG emissions into kgs of CO_2 equivalents (Lal, 2004; Lal et al., 2020). The amount of GHG emissions from different farm operations, such as machinery, livestock, fuel, diesel, electricity, nitrogen (N), phosphorus (P_2O_5), potassium (K_2O), manure, herbicides, pesticides and fungicides, were estimated using the input specific CO_2 emission coefficients (Table 3). The total CO_2 emission was obtained by multiplying the total energy consumed from each input item with the CO_2 emission coefficient.

3 Results and discussion

3.1 Energy consumption and output

The average yield of okra and its corresponding by-products were 8.5 and 6.5 t ha^{-1}, respectively. The tomato yield was 23.7 t ha^{-1} while the by-product yield was 11.5 t ha^{-1}.

Table 3 Coefficients of GHG emission of agricultural inputs

Inputs	Unit	GHG coefficient (kg CO_2eq $unit^{-1}$	References
Machinery	MJ	0.071	Dyer and Desjardins (2006)
Bullocks (pair)	hr	2.59	Alfred Gathorne-Hardy (2016)
Diesel	l	0.94	Fluck (1992)
Gasoline	l	0.85	Fluck (1992)
Electricity	kWh	0.608	Nabavi-Pelesaraei et al. (2014)
Chemical fertilizers			
Nitrogen (N)	kg	1.3	Lal (2004)
Phosphorus (P_2O_5)	kg	0.2	Lal (2004)
Potassium (K_2O)	kg	0.2	Lal (2004)
Farmyard manure	kg	0.126	Pishgar-Komleh et al (2013)
Biocide			
Herbicide	kg	6.3	Lal (2004)
Insecticide	kg	5.1	Lal (2004)
Fungicide	kg	3.9	Lal (2004)

Total energy consumption from different sources for production of okra and tomato was 8828.71 and 4798.66 MJ ha^{-1}, respectively (Table 4).

The energy output of okra was calculated as 184.89% higher than the energy input, while in tomato, it was 658.46% higher. The overall energy requirement for okra cultivation estimated in this study was slightly lower than earlier reported value (Tuti et al., 2012; Sumita & Singh, 2018). Similarly, in tomato, the present study's input energy requirement was less than that reported by other workers, viz. 131,634.19 MJ ha^{-1} (Heidari & Omid, 2011); 53,222 MJ ha^{-1} (Karamian et al., 2021); 104,915.10 MJ ha^{-1} Saltuk (2019); 56,127.77 MJha^{-1} (Yelmen et al., 2019); 116,768.4 MJ ha^{-1} (Taki et al., 2013). In conventional method of cultivation, the yield of okra has been reported as 9.26–9.69 t ha^{-1} (Kandeyang et al., 2010; Singh et al., 2017).The energy output in case of tomato was estimated at 108,000 MJ ha^{-1} (Taki et al., 2013); 50,400 MJ ha^{-1} (Karamian et al., 2021); 128,000 MJ ha^{-1} (Saltuk, 2019); 245,246.78 MJ ha^{-1} (Yelmen et al., 2019).

Contributions of direct and indirect energy sources in total energy consumed in tomato cultivation were 3241MJ ha^{-1} and 1558 MJ ha^{-1}, respectively. In case of okra, these values were 6543 MJ ha^{-1} and 2286 MJ ha^{-1}, respectively. Okra had higher contribution from direct energy (74.1%), whereas indirect energy (32.5%) was major contributor in case of tomato, indicating increased use of non-conventional energy, viz. diesel and fertilizer. Use of both renewable and nonrenewable forms of energy was observed in okra and tomato. The renewable energy component is high in okra, suggesting that the traditionally available energy sources are still major input to farming in the tribal belt. Tomato registered higher nonrenewable energy input (72.3%), suggesting more use of diesel, electricity, and machinery, while okra had the maximum share of renewable energy (59.6%).

3.2 Pattern of energy use

The percentage share (%) and component-wise energy inputs (MJ $unit^{-1}$) involved indifferent farm operations of vegetable cultivation are shown in Fig. 1 and Table 5. The cumulative energy used for okra production in various farm activities was 8828.71 MJ ha^{-1} (Table 5). Human energy contributed the largest share (27.62%), followed by petrol (27.31%), farmyard manure (13.59%), animal energy (13.22%), electricity (6.05%), seed (5.22%), and machinery (5.08%). The animal energy was consumed primarily for land preparation, intercultural practices, and transport of vegetables from the field to the market place. The use of irrigation by pump and electricity was minimal at 1.87%, as agriculture is mostly rain-fed. Energy use in terms of chemicals was negligible (0.03%), as most of the farmers used cow urine and extract neem oil in place of insecticide. Fertilizer was not applied, and farmers mostly depended on FYM derived from cow dung. Tillage operation was done using draught animals, and there was no

Table 4 Yield, input and output energy for summer vegetable under Chotanagpur plateau region

Item	Okra	Tomato
Output main (kg ha^{-1})	8500	23,700
By-products (kg ha^{-1})	6500	11,500
Total output (kg ha^{-1})	15,000	35,200
Energy input (MJ ha^{-1})	8828.71	4798.66
Energy output (MJ ha^{-1})	25,152.50	36,396.00

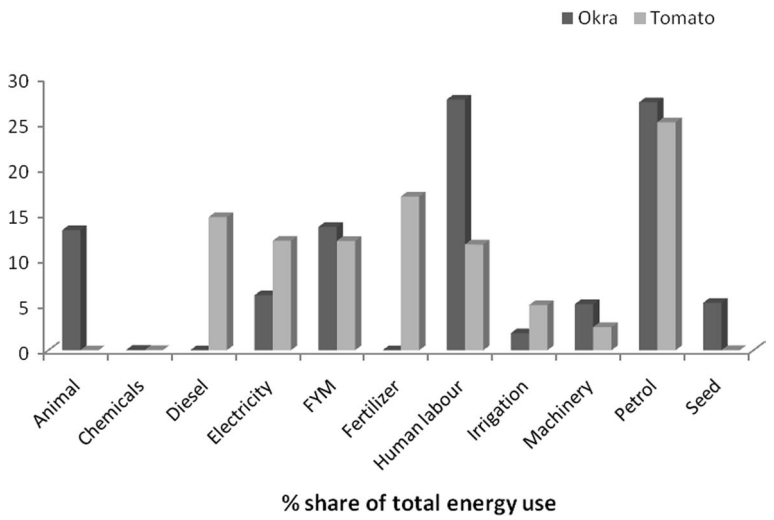


Fig. 1 Percent share of inputs in total energy for okra and tomato production in Chotanagpur plateau region

Table 5 Source wise distribution of input energy (MJ ha^{-1}) for okra and tomato production in Chotanagpur plateau region

Energy input	Okra	Tomato
Animal energy	1167.43 (13.22[a])	0.00 (0.00)
Chemicals	3.00 (0.03)	0.50 (0.01)
Diesel	0.00 (0.00)	703.88 (14.67)
Electrical	533.99 (6.05)	578.49 (12.06)
FYM	1200.00 (13.59)	577.50 (12.03)
Fertilizer	0.00(0.00)	813.00 (16.94)
Human labor	2438.29 (27.62)	558.81 (11.65)
Irrigation	165.40 (1.87)	238.71 (4.97)
Machinery	448.30 (5.08)	121.78 (2.54)
Petrol	2411.50 (27.31)	1205.75 (25.13)
Seed	460.80 (5.22)	0.25 (0.01)
Total (MJ ha^{-1})	8828.71(100.00)	4798.66 (100.00)

[a]Figures on parenthesis indicate the percent share in total energy

use of tractor/other machinery except sprayer and pump set. Petrol consumption was the primary input in tomato production, mainly used for the transportation of produce from field to market place. It contributed around 25.13% of total energy consumption. The energy usage for tomato production in different farm activities was 4798.66 MJ ha^{-1}. The second highest contributor was fertilizer (16.94%), followed by diesel (14.67%) of energy use. The shares of electricity, farmyard manure, human energy, irrigation, and machinery inputs were 12.06, 12.03, 11.65, 4.97, and 2.54%, respectively. Energy use on chemicals and seed had the least share with 0.01%, as traditional methods were used to carried out most management practices. Tuti et al. (2012) stated energy input for okra production as 12,536 MJ ha^{-1} which is 29.57% higher than the value estimated in the present study. The lower value of input energy in the present study can be attributed to minimal use of chemicals, fertilizer, and seed in the study region. Heidari and Omid

(2011) also confirmed that among other input energy sources for tomato production, diesel fuel (49.02%) and fertilizers (24.01%) were major contributors to total energy input.

Direct energy inputs like diesel, electricity, irrigation, human labor, and animal energy contributed 74.1% (6543 MJ ha^{-1}) to the total energy use in tomato production. In contrast, indirect energy inputs like seeds, fertilizer, FYM, plant protection chemicals, and machinery contributed 25.9% (2286 MJ ha^{-1}) for okra production. Similarly, for tomato production, the share of direct energy input was 67.5% (3241 MJ ha^{-1}), while the share of indirect energy input was 32.5% (1558 MJ ha^{-1}). In case of okra, renewable energy inputs like human labor, animal energy, seeds, FYM, and irrigation were 27.7% (1330 MJ ha^{-1}) of total energy use. In contrast, nonrenewable energy inputs like diesel, fertilizer, plant protection chemicals, machinery, and electricity were 72.3% (3468 MJ ha^{-1}) for tomato production. Similarly, for okra production, the renewable energy contribution was 59.6% (5258 MJ ha^{-1}), whereas the share of nonrenewable energy input was 40.4% (3570 MJ ha^{-1}) (Fig. 2). In okra, Taki et al. (2013) reported direct and indirect energy shares as 66% and 34% of the total input energy while the contribution of renewable energy (19%) was smaller as compared to nonrenewable resources (81%). Heidari and Omid (2011) stated that, in Iran, the share of conventional energy for tomato productions was 90%.

3.3 Energy usage in agronomic operations

Harvesting and transportation activities were the top contributors (41.6%) to the total energy consumption in agronomic practices for okra production, whereas in case of tomato cultivation, fertilizer and FYM had the highest energy share (30.97%). In okra, the second highest contributor toward energy input was land preparation (21.24%), followed by

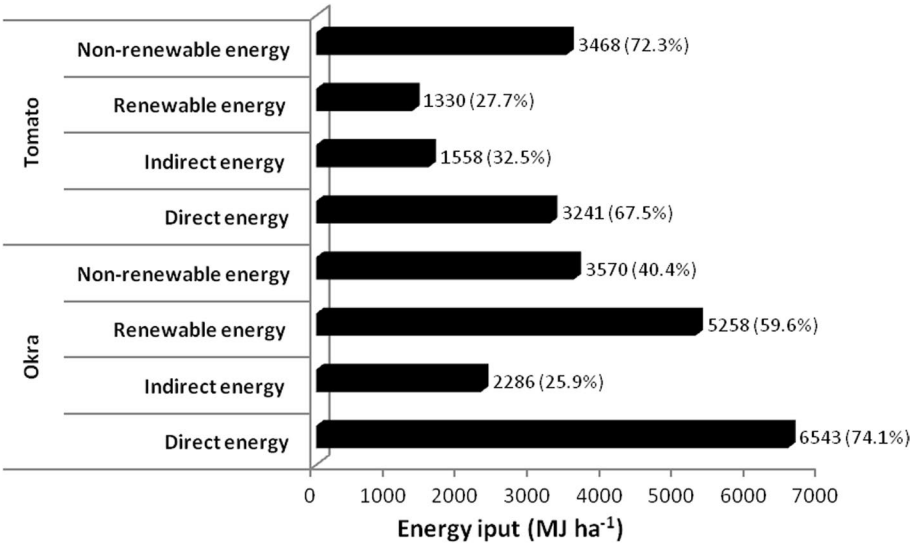


Fig. 2 Source wise distribution of direct, indirect, renewable and non-renewableinput energy for okra and tomato production in Chotanagpur plateau region

fertilizer (16.68%), irrigation (8.77%), seed (8.12%), intercultural operations (3.72%), and plant protection (0.31%). In tomato cultivation, the second highest contributor toward input was harvesting and transportation (29.21%), irrigation (18.5%), land preparation (17.61%), intercultural operations (2.52%), seed (0.92%), and plant protection (0.27%).

3.4 Energy indices

The okra production's energy ratio was estimated to be around 2.85 compared to 7.58 in tomatoes (Table 6), indicating that in Eastern Plateau Region, the tomato production is more energy efficient than okra production. The energy ratio greater than one shows the system's energy gain, while less than 1 represents the system's energy loss (Conforti & Giampietro, 1997). Other studies conducted on tomato indicated energy use efficiency of 1.48 (Heidari & Omid, 2011) and 0.32 (Canakci & Akinci, 2006). Furthermore, net energy gain, energy productivity, and specific energy were 16,323.79 MJ ha^{-1}, 0.56 kg MJ^{-1}, and 1.79 MJ kg^{-1}, respectively, for okra and 31,597.34 MJ ha^{-1}, 1.54 kg MJ^{-1}, and 0.8 MJ kg^{-1}, respectively, in tomato. Energy productivity of tomato was 63.64% higher as compared to okra. On account of higher fuel expenditure on transportation of produce from farm to market place, the specific energy requirement in okra (1.79 MJ kg^{-1}) was higher than that of tomato (0.8 MJ kg^{-1}). Bhargavi and Behera (2019) reported energy productivity was 423.4 and 485.8 g MJ^{-1}, whereas the specific energy was 2.36 and 2.06 for maize-pea-okra and okra-wheat system, respectively.

For greenhouse cultivated tomato in Turkey, the specific energy and energy efficiency for tomato cultivation were 2.49 MJ kg^{-1} and 0.40 kg MJ^{-1}, respectively (Canakci & Akinci, 2006). Saltuk (2019) reported that energy efficiency, specific energy, and productivity were reported as 1.22, 0.66 $MJkg^{-1}$, and 23,084.90 $MJha^{-1}$, respectively, for greenhouse tomato production. Elhag and Boteva (2019) reported the energy use efficiencies of 1.462, 1.198, and 1.147 for greenhouse cultivated tomato–pepper system; tomato–pepper–cucumber, and tomato cucumber systems, respectively. Hasan and Taki (2018) studied energy use in tomato production in Turki and reported cumulative degree of perfection and renewability indicator for tomato production as 1.62 and 0.38, respectively.

In the present study, the energy profitability was 10.32 and 65.13 for okra and tomato cultivation, respectively. Lower the energy profitability, higher is the consumption of human energy, and fuel for transportation of produce. The human energy profitability (HEP) was estimated at10 and 65 for okra and tomato. A low level of mechanization in okra indicates increased human energy input (HEP) by 67%, compared to tomato production. It shows that the region has not witnessed large-scale applications of farm machinery in vegetable production.

It was observed that the farmers followed a mix of standard and traditional agricultural practices in the region. In the process of okra production, most preharvest processes

Table 6 Energy indicators and forms of energy for okra and tomato production in Chotanagpur plateau region

Item	Okra	Tomato
Energy ratio	2.85	7.58
Energy productivity (kg MJ^{-1})	0.56	1.54
Specific energy (MJ kg^{-1})	1.79	0.8
Net energy gain (MJ ha^{-1})	16,323.79	31,597.34

(tillage, seedling, weed control, irrigation, etc.) are of traditional nature. Tomato production is bit mechanized as tillage was done using a tractor. The state's farm power availability is much lower (1.21 kW ha^{-1}during 2016–2017) than all India average (2.03 kWh ha^{-1}). Custom hiring of farm machinery is slowly gaining pace in the region among small and marginal farmers. Organic fertilizers are preferred over inorganic ones. Only 18% of agricultural land had irrigation facilities, and the rest were dependent on monsoon rains. The farmers used animal-drawn plough as well as tractor powered implements for preparing their fields. All the agricultural activities after land preparation were done by human labor using small farm implements and tools, viz. *khurpi*, spade, sickle, knapsack sprayer, etc. If there is no timely precipitation, almost all farmers have drilling kits to pump underground water for irrigation. Harvesting was done manually for all the crops as the farm was small and fragmented and was not suitable for plying large harvesting machinery.

3.5 Economics of okra and tomato production

The per hectare cost in okra cultivation was estimated at Rs. 74,576.15 (Table 7a). The total cost of human labor engaged in various agricultural operations such as field preparation, sowing, earthing up, weeding, irrigation, plant protection activities, harvesting/fruit collection, etc. (Rs. 39,993/-) constituted the major component of (53.63%) of the total cost. The labor cost of okra harvesting/collection turned out to be a significant contributor (28.13%) to total labor cost. In the total labor costs for various agricultural operations, the cost of okra harvesting/collection was significant which contributed 28.13% to total labor cost. Other agricultural operations, i.e., ridge and furrow preparation, sowing with *khurpi*, initial field preparation, weeding, earthing-up, etc., contributed 18.75%, 15.00%, 11.25%, 9.38%, and 8.44%, respectively, of total labor cost. Okra yield was 8.5tha^{-1}, fetching a total return of Rs. 1,27,500/- and net return of Rs. 52,923.8 per hectare with a benefice–cost (BC) ratio of 1.71, indicating a reasonably good return in okra cultivation by the farmers of Jharkhand's plateau region.

Table 7b shows the details of per hectare cost and returns in the cultivation of hybrid tomato in the plateau region of Chotanagpur. The cost of tomato production was Rs. 30,120.66 per hectare, constituting labor costs involved in different agricultural operations as the significant cost (34.79%), followed by seed cost (24.90%). The total labor cost involved expenditure on various agricultural operations such as seedbed preparation, seed sowing, irrigation, uprooting of seedlings and transplanting, fertilizer application, weeding, plant protection, bund maintenance, and product harvesting/collection. Out of these operations, product harvesting and its collection was observed to be the most labor-intensive operation, as it contributed a maximum (26.84%) to the total labor cost, followed by weeding (21.47%), fertilizer application during seedbed preparation, and primary field application (15.20%), bund maintenance (14.32%), and plant transplanting (7.16%).

The yield of hybrid tomato obtained was 23.7 t ha^{-1}, fetching a total return of Rs. 2,37,000/- per hectare and net return of Rs. 2,06,877 per hectare with a BC ratio of 7.87, indicating a reasonably good return in tomato cultivation by the farmers of Jharkhand's plateau region. Selection of high yielding genotypes which are resistant to pests and diseases is important to obtain higher net returns (Singh et al., 2020). Adoption of high yielding hybrids and better market price led to higher net returns from both these crops. Lower BC ratio in present study was mainly due to higher cost of cultivation on account of higher labor requirement in okra.

Table 7 (a) Cost and returns in okra cultivation, (b) cost and returns in tomato cultivation

S. no.	Operations/materials	Input/output	Rate (Rs. unit^{-1})	Cost (Rs.)
(a)				
Costs				
1	Field preparation			
	1.1 Ploughing (02), clod crushing and planking (draught animal-hr)	110	120	13,200.00
	1.2 FYM (q)	40	80	3200.00
	1.3 Labor engaged (man days)	15	300	4500.00
2	Ridge and furrow preparation (man days)	25	300	7500.00
3	Seeds (kg)	9	550	4950.00
	3.1 Sowing with khurpi (man days)	20	300	6000.00
	3.2 Earthing- up (man days)	11.25	300	3375.00
4	Weeding-Manual (man days)	12.5	300	3750.00
5	5.1 Irrigation-4 times	4	200	800.00
	5.2 Labor engaged (man days)	4.56	300	1368.00
6	Plant protection			
	6.1 Neem extract (litre)	12	550	6600.00
	6.2 Labor engaged (man days)	7.5	300	2250.00
7	Harvesting and collection (man days)	37.5	300	11,250.00
8	Total working capital (Rs.)			68,743.00
9	Interest on working capital for 4 months (% year^{-1})		8	1833.15
10	Rental value of own rainfed land for 4 months (Rs ha^{-1} year^{-1})		12,000	4000.00
Total cost (A)			74,576.15	
Returns				
	Yield (kg ha^{-1})	8500		
	Average output market rate (Rs kg^{-1})		15	
Total return (B)			127,500	
	Net return (B-A)			52,923.85
	B:C ratio			1.71

Table 7 (continued)

S. no.	Operations/materials	Input/output	Rate (Rs. unit^{-1})	Cost (Rs.)
(b)				
Costs				
1	Seed bed preparation			
	Ploughings by tractor and planking for leveling (tractor-hr)	0.62	700	434.00
	1.1 Fertilizer/Manure			
	Cow dung manure (q)	1.85	80	148.00
	Potash (kg)	5	11.54	57.70
	Labor engaged (man days)	0.31	300	93.00
2	Seed (kg)hybrid	0.25	30,000	7500
	2.1 Sowing-manually (man days)	0.31	300	93.00
3	Irrigation			
	3.1 Irrigation (01 time)	1	200	200.00
	3.2 Labor engaged (man days)	0.31	300	93.00
4	Up-rooting of seedlings (man days)	0.62	300	186.00
5	Main Field Preparation			
	5.1 Ploughing and planking (tractor- hr)	2.5	700	1750.00
6	Fertilizer application			
	6.1 FYM (q)	17.5	80	1400.00
	6.2 Urea (kg)	12.5	6	75.00
	6.3 Labor charges (man days)	5	300	1500.00
7	Transplanting			
	7.1 Labor engaged (man days)	2.5	300	750.00
8	Irrigation (6 times) 216 m^3	6	200	1200.00
9	Weeding			
	9.1 Labor charges (man days)	7.5	300	2250.00

Table 7 (continued)

S. no.	Operations/materials	Input/output	Rate (Rs. unit^{-1})	Cost (Rs.)
10	Plant protection			
	10.1 Neem extract (litre)	4	550	2200.00
	10.2 Labor engaged (man days)	4	300	1200.00
11	Bund maintenance/ support of bamboo sticks/ partial harvesting (man days)	5	300	1500.00
12	Harvesting and collection (man days)	9.37	300	2812.50
13	Total working capital (Rs.)			25,442.20
14	Interest on working capital for 4 months (% year^{-1})		8	678.46
15	Rental value of own rainfed land for 4 months (Rs ha^{-1} year^{-1})		12,000	4000.00
Total cost (A)			30,120.66	
Returns				
	Yield (kg ha^{-1})	23,700		
	Average output market rate (Rs kg^{-1})		10	
Total return (B)			237,000.00	
	Net return (B–A)		206,879.34	
	B:C ratio		7.87	

3.6 GHG estimation

The GHG emissions from inputs used in the cultivation of okra were estimated at 875.41 kg CO_2eq ha^{-1} (Fig. 3). FYM with 504 kg CO_2eq emissions per hectare was the main contributor (57.57%) to total GHG emissions from okra. Bullock energy was second largest contributor (34.20%), followed by petrol (4.85%), electricity (3.11%), insecticide (0.17%), and machinery (0.09%). The chemical fertilizer was not used in okra cultivation.

Total GHG emission from tomato production was 322.75 kg CO_2eq ha^{-1}. In this crop also, the FYM had maximum share (75.15%) in total GHG emissions followed by electricity (9.13%), petrol (6.58%), nitrogen (5.03%), and diesel (3.64%), whereas shares of emission from potassium (0.31%), machinery (0.07%), and insecticide (0.08%) were negligible. The total amount of CO_2 emissions from greenhouse tomato production was recorded by Taki et al. (2013) as 4280.849 kg CO_2eq ha^{-1}. Range of values of GWP has been reported in the literature. In Spain, the GWP of producing 1 ton of tomato was 467.44 kg CO_2eq (Bosona & Gebresenbet, 2018) while in Iran it was 65.8 kg CO_2eq (Zarei et al., 2019). Demirtas (2021) determined the energy use in the production of one kg paste tomato as 175 kJ and CO_2 emission as 15.88 g CO_2 in Bursa region of Turkey. Persiani et al. (2021) reported the GHG emission of 2312 kg CO_2 eq ha^{-1} $year^{-1}$ for tomato production while in Turkey it was close to 1724.70 kg CO_2 ha^{-1} (Kulekci & Sari, 2020). Ronga et al. (2019) investigated the environmental impact of tomato production and reported GWP of the organic cropping system as 67.49 kg CO_2eq t^{-1} and 55.16 kg CO_2eq t^{-1} in the conventional cropping system. The annual energy requirements varied

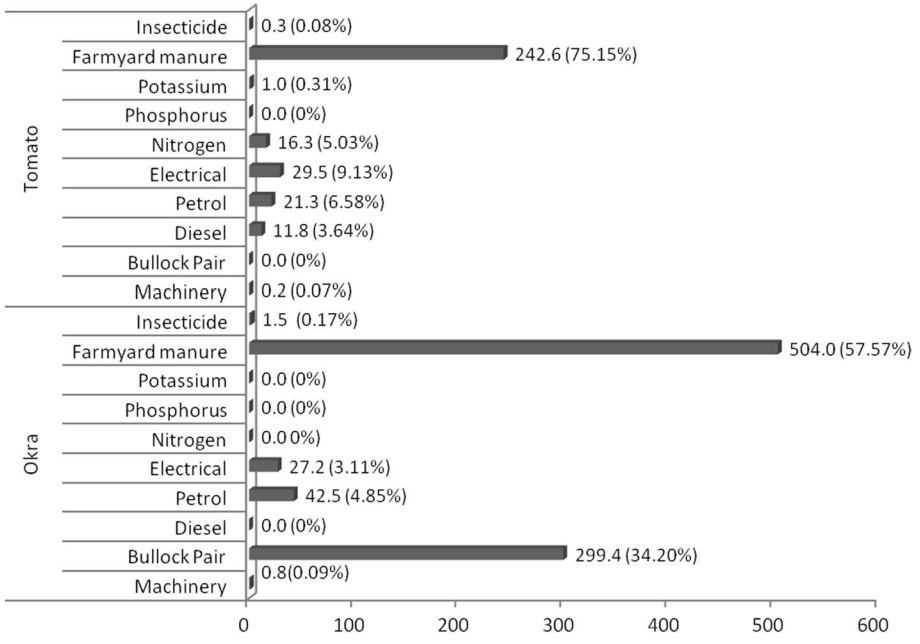


Fig. 3 Greenhouse gas (GHG) emissions (kg CO_2eq ha^{-1}) from different inputs used in okra and tomato production in Chotanagpur plateau region

between 0.8 and 160 MJ kg^{-1}, and GHG emissions from tomato were between 0.1 and 10.1 CO_2eq kg^{-1} for seven different tomato cultivation systems in southern and central Europe (Ntinas et al., 2017).

Although the available literature on GHG emission from tomato and okra is scanty, substantial research has been carried out in other vegetables. In cucumber grown under the greenhouse, the overall GHG emissions were estimated as 82.724 kg CO_2eq ha^{-1} (Pishgar-Komleh et al., 2013) while that for general vegetable production was 17,068.37 kg CO_2eq ha^{-1} (Nourani and Bencheikh, 2019). In case potato cultivation, the estimated CO_2 emissions were 993 kg CO_2eq ha^{-1} (Pishgar-Komleh et al., 2012).

4 Conclusions

The present study analyzed the energy input and output patterns of okra and tomato crops cultivated in the eastern region of India. Analysis indicated that okra was most energy-intensive crop while tomato provided more energy output leading to higher energy ratio (7.58). Okra cultivation consumed 1.84 times higher energy as compared to tomato crop while the total energy output from tomato crop was 1.45 time higher over okra. Tomato cultivation had lower specific energy consumption (0.80 MJ kg^{-1}) and higher energy productivity (1.54 kg MJ^{-1}) as compared to 1.79 MJ kg^{-1} and 0.56 kg MJ^{-1} observed in case of okra. Higher specific energy in okra implied greater use of input energy in terms of labor, fuel, FYM, and animal energy. As compared to okra, the tomato crop produced 93% higher net energy gains per unit area.

In terms of physical productivity, the tomato was more productive with an average yield of 23.7 t ha^{-1} as compared to 8.5 t ha^{-1} observed in okra. From the economic perspective, okra is a more capital-intensive crop as it required working capital of Rs. 68,743 over the entire growing season. Tomato cultivation was more profitable system with the benefit–cost ratio of 7.87 as against 1.71 observed in okra. More involvement of human labor, higher cost of cultivation and lower physical productivity make okra less remunerative than tomato. Lower shares of chemical fuels in inputs imply lower level of mechanization of farm operations in the region. As compared to tomato, okra crop had 2.71 times higher carbon footprints due to its input intensiveness. Analysis showed that tomato cultivation was more environment friendly as it had lower GHG emissions (322.75 kg CO_2eq ha^{-1}) as compared to tomato (875.41 kg CO_2eq ha^{-1}). FYM was the largest contributor to total GHG emission in okra (57.57%) and tomato (75.15%). Under the same environmental condition, tomato production is more favorable as it produces lower carbon footprint than okra.

Given the prevalence of conventional tillage practices in this region, both tomato and okra required lower energy inputs and higher energy productivity. Comparatively, tomato crop having higher system productivity, lower energy input requirement, and higher energy ratio may be a better option for wide-scale adoption in agroecosystem of eastern plateau region of India. Under the global climate change scenario, tomato being less energy intensive, having lower CO_2 emissions and better environmental and economic gains over okra, should be preferred for cultivation in plateau region of India. Outcomes of our study could be useful to policymakers for framing policies for summer season cultivation of vegetable crops.

Acknowledgements The authors are highly thankful to ICAR-Research Complex for Eastern Region, Patna-800014 (India), and Centre for Energy and Environment, Delhi Technological University, Delhi-110042 (India), for providing amenities to complete this study.

Declarations

Conflict of interest Funding for the present study was authorized by the Indian Council of Agricultural Research, New Delhi. The authors, employed by ICAR-Research Complex for Eastern Region, do not declare any conflict of interest concerning the information discussed in the present given study. Its contents, including any views and conclusions shared, are those of the authors themselves and do not generally constitute the Indian Council of Agricultural Research's policy, New Delhi.

References

Bhargavi, B., & Behera, U. (2019). System productivity and energetics of high-value crops embedded diversified cropping systems. *International Journal of Current Microbiology and Applied Sciences, 8*, 1895–1905.

Bosona, T., & Gebresenbet, G. (2018). Life cycle analysis of organic tomato production and supply in Sweden. *Journal of Cleaner Production, 196*, 635–643.

Canakci, M., & Akinci, I. (2006). Energy use pattern analyses of greenhouse vegetable production. *Energy, 31*, 1243–1256.

Census. (2011). Office of the Registrar General & Census Commissioner, Ministry of Home Affairs, Government of India. Retrieved January 5, 2019, from https://www.censusindia.co.in/district/ramgarh-districtjharkhand-361

Conforti, P., & Giampietro, M. (1997). Fossil energy use in agriculture: An international comparison. *Agriculture, Ecosystems and Environment, 65*(3), 231–243.

Demirtas, M. (2021). Evaluation of energy use and carbon dioxide emissions from the consumption of fossil fuels and agricultural chemicals for paste tomato cultivation in the Bursa region of Turkey. *Environmental Science and Pollution Research, 28*(21), 27132–27146.

Devasenapathy, P., Senthilkumar, G., & Shanmugam, P. M. (2009). Energy management in crop production. *Indian Journal of Agronomy, 54*(1), 80–90.

Dyer, J. A., & Desjardins, R. L. (2006). Carbon dioxide emissions associated with the manufacturing of tractors and farm machinery in Canada. *Biosystems Engineering, 93*(1), 107–118.

Elhag, M., & Boteva, S. (2019). Conceptual assessment of energy input-output analysis and data envelopment analysis of greenhouse crops in Crete Island, Greece. *Environmental Science and Pollution Research, 26*(35), 35377–35386.

Esengun, K., Erdal, G., Gunduz, O., & Erdal, H. (2007). An economic analysis and energy use in stake-tomato production in Tokat province of Turkey. *Renewable Energy, 32*, 1873–1881.

Fluck, R. C. (1992). Energy in farm production. In R. C. Fluck (Ed.), *Energy in world agriculture* (6th ed., pp. 218–267). Elsevier.

Gathorne-Hardy, A. (2016). The sustainability of changes in agricultural technology: The carbon, economic and labour implications of mechanisation and synthetic fertiliser use. *Ambio, 45*, 885–894.

Ghorbani, R., Mondani, F., Amirmoradi, S., Feizi, H., Khorramdel, S., Teimouri, M., Sanjani, S., Anvarkhah, S., & Aghel, H. (2011). A case study of energy use and economic analysis of irrigated and dryland wheat production systems. *Applied Energy, 88*, 283–288.

Hamedani, S. R., Shabani, Z., & Rafiee, S. (2011). Energy inputs and crop yield relationship in potato production in Hamadan province of Iran. *Energy, 36*, 2367–2371.

Hasan, Y., & Taki, M. (2018). Assessment of tomato production process by cumulative exergy consumption approach in greenhouse and pen field conditions: Case study of Turkey. *Energy, 156*, 401–408.

Heidari, M. D., & Omid, M. (2011). Energy use patterns and econometric models of major greenhouse vegetable productions in Iran. *Energy, 36*, 220–225.

Huang, B., Shi, X., Yu, D., Oborn, I., Blomback, K., & Pagella, T. F. (2006). Environmental assessment of small-scale vegetable farming systems in peri-urban areas of the Yangtze River Delta Region, China. *Agriculture, Ecosystems and Environment, 112*, 391–402.

Ju, X. T., Kou, C. L., Christie, P., Dou, Z. X., & Zhang, F. S. (2007). Changes in the soil environment from excessive application of fertilizers and manures to two contrasting intensive cropping systems on the North China Plain. *Environmental Pollution, 145*, 497–506.

Kandeyang, S., Jerai, M. C., & Pandey, A. C. (2010). Effect of cropping system and fertilizer levels on productivity and economics of okra (*Abelmoschus esculentus* L. Moench). *International Journal of Tropical Agriculture, 28*(1–2), 287–289.

Karamian, F., Mirakzadeh, A. A., & Azari, A. (2021). The water-energy-food nexus in farming: Managerial insights for a more efficient consumption of agricultural inputs. *Sustainable Production and Consumption, 27*, 1357–1371.

Kastner, T., Rivas, M. J. I., Koch, W., & Nonhebel, S. (2012). Global changes in diets and the consequences for land requirements for food. *Proceedings of the National Academy of Sciences, 109*, 6868–6872.

Khoshnevisan, B., Rafiee, S., Omid, M., & Mousazadeh, H. (2013). Applying data envelopment analysis approach to improve energy efficiency and reduce GHG (greenhouse gas) emission of wheat production. *Energy, 58*, 588–593.

Khoshroo, A., Izadikhah, M., & Emrouznejad, A. (2021). Energy efficiency and congestion considering data envelopment analysis and bounded adjusted measure: A case of tomato production. *Journal of Cleaner Production, 328*, 129639.

Kulekci, M., & Sari, M. M. (2020). Reduction of the greenhouse gas emission and energy optimization for tomato production. *Fresenius Environmental Bulletin, 29*(7), 6168–6180.

Kumar, M., Kumar, R., Rangnamei, K. L., Das, A., Meena, K. L., & Rajkhowa, D. J. (2019). Crop diversification for enhancing the productivity for food and nutritional security under the Eastern Himalayas. *Indian Journal of Agricultural Sciences, 89*(7), 1157–1161.

Lal, B., Gautam, P., Panda, B. B., Tripathi, R., Shahid, M., Bihari, P., Guru, P. K., Singh, T., Meena, R. L., & Nayak, A. K. (2020). Identification of energy and carbon efficient cropping system for ecological sustainability of rice fallow. *Ecological Indicators, 115*, 106431.

Lal, R. (2004). Carbon emission from farm operations. *Environment International, 30*, 981–990.

Liu, R., Zhang, P., Wang, X., Chen, Y., & Shen, Z. (2013). Assessment of effects of best management practices on agricultural non-point source pollution in Xiang xi River watershed. *Agricultural Water Management, 117*, 9–11.

Mittal, V. K., Mittal, J. P., & Dhawan, K. C. (1985). Research digest on energy requirements in agricultural sector. Coordinating cell, AICRP on energy requirements in agricultural sector. Punjab Agricultural University, Ludhiana.

NABARD. (2017). State focus paper 2016-17:1–14 Jharkhand. https://www.nabard.org

Nabavi-Pelesaraei, A., Abdi, R., Rafiee, S., & Mobtaker, H. G. (2014). Optimization of energy required and greenhouse gas emissions analysis for orange producers using data envelopment analysis approach. *Journal of Cleaner Production, 65*, 311–317.

Nassiri, S. M., & Singh, S. (2009). Study on energy use efficiency for paddy crop using data envelopment analysis (DEA) technique. *Applied Energy, 86*(7–8), 1320–1325.

Nourani, A., & Bencheikh, A. (2019). Energy use pattern and greenhouse gas emission in systems for greenhouse vegetable production. *Agricultural Science and Technology., 11*(3), 240–246.

Ntinas, G. K., Neumair, M., Tsadilas, C. D., & Meyer, J. (2017). Carbon footprint and cumulative energy demand of greenhouse and open-field tomato cultivation systems under Southern and Central European climatic conditions. *Journal of Cleaner Production, 142*, 3617–3626.

Ozturk, H. H., Kücukerdem, K., Mutlu, N., Gozubuyuk, Z., & Atay, U. (2018). Sustainability indicators for energy use and environmental impacts in agricultural production. *Journal of Agricultural Machinery Science, 14*(3), 199–204.

Patra, A. K., Garnayak, L. M., Behera, B. B., Swain, D., & Paikaray, R. K. (2021). Productivity, resource use efficiency and economics of rice (Oryza sativa)-based bio-intensive cropping systems in western Odisha. *Indian Journal of Agronomy, 66*(2), 157–162.

Persiani, A., Montemurro, F., Fiore, A., Scazzarriello, R., & Diacono, M. (2021). On-farm fertilizing materials in organic horticulture: Agronomic performance, energy use and GHG emission evaluation. *Archives of Agronomy and Soil Science, 67*(14), 1944–1960.

Pimentel, D., & Pimentel, M. (2008). *Food, energy, and society*. CRC Press.

Pishgar-Komleh, S. H., Ghahderijani, M., & Sefeedpari, P. (2012). Energy consumption and CO_2 emissions analysis of potato production based on different farm size levels in Iran. *Journal of Cleaner Production, 33*, 183–191.

Pishgar-Komleh, S. H., Omid, M., & Heidari, M. D. (2013). On the study of energy use and GHG (greenhouse gas) emissions in greenhouse cucumber production in Yazd province. *Energy, 56*, 63–71.

Ronga, D., Gallingani, T., Zaccardelli, M., Perrone, D., Francia, E., Milc, J., & Pecchioni, N. (2019). Carbon footprint and energetic analysis of tomato production in the organic vs the conventional cropping systems in Southern Italy. *Journal of Cleaner Production, 220*, 836–845.

Saini, K. S., Kaur, J., Singh, T., Chahal, T. S., & Manhas, S. S. (2020). Productivity, economics and energetics of potato (*Solanum tuberosum*)-based systems as an alternative to rice (*Oryza sativa*)-wheat (*Triticum aestivum*) in Doaba region of Punjab. *Indian Journal of Agronomy, 65*(1), 125–129.

Saltuk, B. (2019). Energy efficiency of greenhouse tomato production in Turkey: A case of siirt province. *Fresenius Environmental Bulletin, 28*(8), 6352–6357.

Saxena, M., & Gandhi, C. P. (2015). All India 2013–14 (final estimates of Tomato), Department of Agriculture & Cooperation, Indian Horticulture Database-2014. National Horticulture Board, Ministry of Agriculture, Government of India (p. 181).
Singh, S. K., Yadav, R. B., Singh, J., & Singh, B. (2017). Organic farming in vegetables. IIVR technical bulletin no. 77, ICAR-IIVR, Varanasi (p. 47).
Singh, A. K., Bikash, D., Mali, S. S., Bhavana, P., Shinde, R., & Bhatt, B. P. (2020). Intensification of rice-fallow cropping systems in the Eastern Plateau region of India: Diversifying cropping systems and climate risk mitigation. *Climate and Development, 12*(9), 791–800. https://doi.org/10.1080/17565529.2019.1696735
Streimikiene, D., Klevas, V., & Bubeliene, J. (2007). Use of EU structural funds for sustainable energy development in new EU member states. *Renewable and Sustainable Energy Reviews, 116*, 1167–1187.
Sumita, P., & Singh, H. B. (2018). Energy inputs and yield relationship in greenhouse okra production by bio-priming. *International Journal of Agriculture, Environment and Biotechnology, 11*(5), 741–746.
Taki, M., Abdi, R., Akbarpour, M., & Mobtaker, H. G. (2013). Energy inputs- yield relationship and sensitivity analysis for tomato greenhouse production in Iran. *Agricultural Engineering International: CIGR Journal, 15*(1), 59–67.
Taskın, O., & Vardar, A. (2016). Use of some renewable energy resources in agricultural production. *Journal of Agricultural Faculty of Uludag University, 30*(1), 179–184.
Toan, P. V., Sebesvari, Z., Bläsing, M., Rosendahl, I., & Renaud, F. G. (2013). Pesticide management and their residues in sediments and surface and drinking water in the Mekong Delta, Vietnam. *Science of Total Environment, 452–453*, 28–39.
Tuti, M. D., Prakash, V., Pandey, B. M., Bhattacharyya, R., Mahanta, D., Bisht, J. K., Kumar, M., Mina, B. L., Kumar, N., Bhatt, J. C., & Srivastva, A. K. (2012). Energy budgeting of colocasia-based cropping systems in the Indian sub-Himalayas. *Energy, 45*, 986–993.
Vanitha, S. M., Chaurasia, S. N. S., Singh, P. M., & Naik, P. S. (2013b). Vegetable statistics, technical bulletin no. 51, IIVR (p. 74).
Vanitha, S. M., Chaurasia, S. N. S., Singh, P. M., & Naik, P. S. (2013a). Vegetable statistics, technical bulletin no. 51, IIVR (p. 118).
Wang, Z., Zhang, X., & Mu, Y. (2008). Effects of rare-earth fertilizers on the emission of nitrous oxide from agricultural soils in China. *Atmospheric Environment, 42*, 3882–3887.
Yelmen, B., Şahin, H. H., & Cakir, M. T. (2019). Energy efficiency and economic analysis in tomato production: A case study of Mersin province in the Mediterranean region. *Applied Ecology and Environmental Research, 17*(4), 7371–7379.
Zarei, M. J., Kazemi, N., & Marzban, A. (2019). Life cycle environmental impacts of cucumber and tomato production in open-field and greenhouse. *Journal of the Saudi Society of Agricultural Sciences, 18*(3), 249–255.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

B. Sarkar[1] · B. Das[1] · P. K. Sundaram[1] · S. S. Mali[1] · A. P. Anurag[1] · A. Upadhyaya[1] · N. Chandra[1] · B. P. Bhatt[2] · A. Kumar[3]

B. Das
bikash41271@gmail.com

P. K. Sundaram
prem.k.sundaram@gmail.com

S. S. Mali
santosh.icar@gmail.com

A. P. Anurag
prakash.anurag121@gmail.com

A. Upadhyaya
aupadhyaya66@gmail.com

N. Chandra
nchandra_09@rediffmail.com

B. P. Bhatt
bhattbp.65@gmail.com

A. Kumar
anilkumar76@dtu.ac.in

[1] ICAR-Research Complex for Eastern Region, Patna 800 014, India

[2] NRM Division, KAB-II, New Delhi, India

[3] Centre for Energy and Environment, Delhi Technological University, New Delhi, Delhi 110 042, India

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1352-2752.htm

Exploring dimensions of the customer experience at budget hotels during the COVID-19 pandemic: a netnography approach

Received 14 March 2022
Revised 19 May 2022
Accepted 14 July 2022

Smitha Girija
Department of Marketing, School of Business, GITAM (deemed to be) University, Visakhapatnam, India

Devika Rani Sharma
Department of Marketing, School of Business, GITAM (deemed to be) University, Hyderabad, India, and

Vaishali Kaushal
Delhi School of Management, Delhi Technological University, Delhi, India

Abstract

Purpose – In 2020, the world encountered travel restrictions because of pandemic, and the hospitality sector across the globe was one of the most affected industries. The purpose of this study is to further explore real-time experiences of guests who stayed in budget hotels and how pandemic has changed the expectations of its customer segment.

Design/methodology/approach – The current study uses netnography to examine customer experience of guests who stayed at budget hotels during the COVID-19 pandemic. A thematic analysis was conducted using NVivo 12 plus on 1,391 customer reviews collected from various travel portals.

Findings – The results suggested personalization and hygiene were the most significant themes that influence customer experience during the COVID-19 pandemic. The authors' analysis revealed that the use of self-service technologies like contactless services during check-in, checkouts and use of Quick Response codes in restaurants and other services would help reduce the perceived risk of guests and enhance the overall customer experience.

Originality/value – The branded budget hotel sector has pumped in lot of money envisaging growth just before the COVID crises emerged. Therefore, the fast recovery is of paramount importance for the sector. Additionally, majority of the pandemic-related studies in hospitality sector have focused on luxury hotels and failed to address the real-time experiences of customers with respect to budget hotels of a developing country. The outcome of this study will be relevant for the budget hotels and policymakers because they face the crucial task of reviving and sustaining enterprises and the industry at large.

Keywords Customer experience, Netnography, Thematic analysis, Personalization, Hygiene, Budget hotels

Paper type Research paper

Introduction

The exponential growth of COVID-19 cases in early 2020 led to a remarkable shrinkage in economic activities due to various mobility restrictions instituted by health authorities. Many hospitality businesses were temporarily closed due to lockdowns, social distancing

Qualitative Market Research: An International Journal
© Emerald Publishing Limited
1352-2752
DOI 10.1108/QMR-03-2022-0039

norms, quarantine and travel bans initiated by authorities to reduce the spread of COVID-19 (Bartik *et al.*, 2020). The tourism and hospitality industry depends heavily on the free movement of people, but due to the COVID-19 pandemic, travel restrictions were being observed on national and international levels. Strict government guidelines to control the spread of the disease brought about extreme challenges for the tourism and hospitality sectors (Gössling *et al.*, 2020a, 2020b). The hotel industry was one of the most negatively impacted economic sectors due to its own nature and business model based on patterns of visitation (Anguera-Torrell *et al.*, 2020). Hotels were potential hotspots for disease spread due to the diversity of populations that visit hotels and other rented establishments, which are enclosed spaces (Gyódi, 2021). Due to the pandemic, hotels have been forced to deal with new kinds of customers who are fearful, economically challenged and under high risk when traveling (Awan *et al.*, 2020).

As per the report on Indian tourism statistics by the Ministry of Tourism (Government of India), more than ten million foreign tourists visited India in 2018–2019, resulting in an enormous contribution to the Indian economy. However, during the pandemic and first lockdown imposed in March 2020 in India, the number of international tourists declined sharply by 74% (UN report 2021). The impact of decline in tourists visit has been devastating for many countries due to revenue loss, and the case of India is no different because 70% of hospitality enterprises are small businesses in India. While the large chains in India's hotel industry have reimagined hospitality during the COVID-19 pandemic, the budget hotels in this sector are still struggling to find the means to survive. Budget hotels or economy hotels may be a branded chain of hotels that have 0–3 stars and a capacity of 50 rooms (Nash *et al.*, 2006; Peng *et al.*, 2015; Ruetz and Marvel, 2011) along with a standardized look with limited food and beverage services (Ruetz and Marvel, 2011). Customers who choose budget or economy hotels are considered to have different lifestyles and consumption patterns than those who choose premium or full-service hotels (Fiorentino, 1995).

Customer experience is considered as pivotal in hospitality sector (Kim and Ritchie, 2014), and it involves the subjective response that customers have to any interactions with a company across multiple touchpoints throughout their purchase journey (Lemon and Verhoef, 2016; Schmitt *et al.*, 2015). Customer experience design has gained prominence due to the ongoing pandemic, because hotels need to not only ensure customers' safety but also prevent the spread of viruses, as hotels fall under high contact services (Bonfanti *et al.*, 2021). The major challenges from the perspective of customer experience in hotel industry during the pandemic are social distancing, concerns on safety and lack of trust (Bharwani and Mathews, 2021). Attempts are made by many hotels to dispel the concerns by redesigning the customer experience journey as this is the only direction toward recovery.

Many studies have explored different aspects of providing safe customer experience to the customers of hospitality sector during the pandemic. Various dimensions such as use of digital platforms for booking (Escandon-Barbosa and Salas-Paramo, 2021); contactless hospitality services (Chen *et al.*, 2021; Bonfanti *et al.*, 2021; Ho and Chen, 2021); hospitality operations – hygiene and cleanliness (Pillai *et al.*, 2021); and guest interaction vs technology (Davari *et al.*, 2022) have been studied in the hospitality industry during the pandemic. However, none of the above studies have addressed the overall customer experience with budget hotels from the perspective of Indian consumers. Because India is one of the most affected countries during pandemic and the contribution of tourism industry to the overall growth of economy is very high, the need to understand the expectations versus delivery gap of the customers who visited budget hotels during this pandemic was quite relevant. Further, the branded budget hotel sector has pumped in lots of money envisaging growth

just before the COVID crises emerged. Therefore, the ramification of this meltdown and further recovery is of immense importance for the Indian academicians and hoteliers. Our objective is to further explore how COVID-19 pandemic has changed the expectations of budget hotels' customer segment.

With the advancement of internet usage, customer online reviews have become an instant way of measuring service performance either good or bad. Such reviews have also provided an opportunity for service companies to respond to consumer call for positive or negative experience (Díaz and Rodríguez, 2018). Many scholars have also used such online reviews for analysis because it gives data on real-time experience. Gunasekar *et al.* (2021) examined the dimensions of wine-experienscape that influence customer satisfaction using online reviews, Kaushal and Yadav (2020) used TripAdvisor customer reviews to explore customer experience of culinary tourism, Philander and Zhong (2016) made use of customer tweets to conduct sentiment analysis of various hotels. Therefore, the current study is aimed to assess the customer experience during the pandemic by using netnographic research on customer reviews of the hospitality provided by budget hotels during the pandemic.

The outcome of this study will be relevant for the budget hotels and policymakers as they face the crucial task of reviving and sustaining enterprises and the industry at large. The paper addresses the following research question: What aspects of the customer experience were relevant to users of budget hotels in India during the pandemic?

Literature review

Understanding customer experience

As today's world is driven by the experiential economy, customer experience is the driving force of the hospitality sector which allows businesses to gain a competitive advantage (Wong, 2013). Offering a flawless customer experience has become a priority to achieve customer loyalty, gain a competitive advantage, create brand equity and generate co-creation value (So and King, 2010; Han and Hwang, 2015). The concept of the customer experience emerged after conventional marketing activities failed to satisfactorily meet customer requirements (Schmitt, 1999). Various factors, such as the emergence of technology, competition, value-added services and superior brands, have contributed to the study of the customer experience (Klaus and Maklan, 2007; Kuo *et al.*, 2009; Ren *et al.*, 2016), and hence, there have been an array of definitions proposed. Few most relevant definitions from the perspective of our study have been reproduced here. Gupta and Vajic (2000)in their study stated experience as an outcome of certain level of interaction that a customer has with different elements created by the service provider, which may even be knowledge or information sharing, whereas Larsen (2007) defined experience from a tourist perspective and related it to the lasting memory a tourist would cherish. Godovykh and Tasci (2020) proposed a more comprehensive definition, defining experience as the totality of negative to positive cognitive, affective, sensory and conative responses evoked by all stimuli encountered in the pre, during and post phases of consumption, filtered through personal differences of consumers, eventually resulting in differential outcomes related to consumers and brands.

Dimensions of customer experience in hotel industry

Studies focusing on the customer experience at hotels have been extensively focused on luxury hotels, full-service hotels, theme-based hotels and other experiential services (Cetin and Walls, 2016; Peng *et al.*, 2015; Walls *et al.*, 2011). Few relevant studies conducted in the context of budget hotels have explored its critical success factors (Brotherton, 2004), which influence customer decisions (Nash *et al.*, 2006; Brunner-Sperdin *et al.*, 2012; Peng *et al.*,

2015). Some existing studies have also tried to explore few dimensions and attributes of budget hotels, such as physical factors (Ren *et al.*, 2016; Luo and Yang, 2016; Peng *et al.*, 2015), human factors (Rahimi and Kozak, 2017), location, brand image (Tanford *et al.*, 2012) and food and beverage services (Ruetz and Marvel, 2011).

The authors conducted a detailed review of studies published in the last 10 years which explored various dimensions of customer experience in the hotel industry to gain more understanding of the topic. The key words used to make literature search are "service experience," "budget hotels," "hotels" and "customer experience." Only the articles pertaining to customer experience in hotel industry are considered for representation in review of literature. Few important studies are mentioned in the Table 1.

From the analysis of the literature, it emerged that researchers have examined the importance of various aspects of the customer experience that covers both physical environment (Han and Hyun, 2017; Ren *et al.*, 2016) and human interactions (Lee *et al.*, 2005; Chittiprolu *et al.*, 2021)in the hospitality industry. The physical environment or mechanics in a hospitality sector include sounds, tastes, landscapes, smells, atmospheres, hygiene and cleanliness (Yuan and Wu, 2008). Many researchers have suggested the importance of safety protocols, cleanliness and hygiene to be adopted in hotels to attract consumers during the ongoing pandemic (Chan *et al.*, 2020; Pillai *et al.*, 2021; Hao *et al.*, 2020). The important role played by technology to contain spread of infection in the ongoing pandemic was also highlighted extensively (Hao *et al.*, 2020; Davari and Jang 2021). Human interaction play an important role in improving the customer experience, as their active participation is vital for offering a superior service (Zeithaml *et al.*, 1985; Jüttner *et al.*, 2013; Cetin and Walls, 2015; Davari *et al.*, 2021). Therefore, human interaction may be understood as a value addition to the physical environment in a service setting. The outbreak of the COVID-19 pandemic has put restrictions on personal interactions due to its contagious nature (Chan *et al.*, 2020), which has also accelerated the use of technology.

Budget hotels have limited resources (Ren *et al.*, 2016) and in such a crisis the need to prioritize and plan for infrastructural improvements is vital to gain the trust of their customers. Hence, it is extremely important to understand the perspective of their guests on the crucial role of various dimensions of physical environment complementing with human interactions wherever unavoidable.

Methodology

Netnography

The internet has prompted researchers to develop new study methodologies that answer many unique questions, such as who, what, when and where amongst various online environments (Turpo Gebera, 2008). Netnography allows researchers to study natural customer activities, such as product reviews or information searches (Kozinets, 2015), by examining consumer behavior in diverse cultures and online communities (Ji *et al.*, 2016). A researcher can use a netnographic approach to understand and identify consumer demands that drive online decision-making by using publicly available information (Kozinets, 2015). Netnography provides many advantages, including low cost, timely compilation of different notions, convenience to study participants, high speculative quality of online opinions and acquisition and the acquisition of meaningful and fascinating results with a small number of messages, if they contain lengthy explanations (Kozinets, 2015). Other research methodologies, including interviews (Walther and Sandlin, 2011), ethnography (McGrath *et al.*, 2013; Li *et al.*, 2021) and even surveys, are equally compatible with netnography (Adjei *et al.*, 2009).

Table 1. Prior studies on customer experience in hotels

Author(s)	Constructs	Outcome
Hao and Chon (2021)	Customer equity, customer trust, delight, experience and customer satisfaction	This study provides practical evidence for hospitality practitioners to consider contactless service in creating memorable experiences, improve customer satisfaction, build trust and add value to hospitality brands
Shen *et al.* (2022)	Acceptance of AR/VR using TAM framework	perceived usefulness, hedonic motivation and price value are important predicting factors for Chinese students' adoption and use of these applications
Zhang *et al.* (2021)	Information search, perceived risks, risk reduction strategies	The outcome revealed that customers face the same kind of risk, they adopt different strategies during the COVID-19 pandemic
Davari and Jang (2021)	Derived six main themes from content analysis: purpose of visit, COVID safety concerns, technology adoption, COVID limitations, exceeded expectation and hospitality of staff. Balance theory approach	The outcome of the study suggests that high-touch and high-tech can indeed co-exist, and that these distinct service delivery modes do not have to be mutually exclusive
Chittiprolu *et al.* (2021)	Satisfiers and dis-satisfiers are identified on the basis of customer reviews using text mining and Herzberg's two-factor theory	This study used Herzberg two-factor theory and found satisfaction and dissatisfaction determinants of customer experience among heritage hotels. The outcome of the study suggests that factors of satisfaction and dissatisfaction remain same, namely, physical ambience, staff behavior, food and value for money
Liu *et al.* (2022)	The study explored the role of SSTs in enhancing customer experience, also assess if SST supplement human performance or replace human performance	The study was two-fold understanding technology both from customers and employee perspective. Customers seemed to have positive response with technology interface
Godovykh and Tasci (2020)	Emotional experience; cognition; conative and sensorial	This study examines empirical and conceptual literature on experience, presents an experience model with four primary components: emotional, cognitive, sensory and conative, and provides a holistic description of experience
Wu and Gao (2019)	Study identified sources of Emotion CX leading to five emotion CX typology having an impact on co-creation in hotel context	The study proposes ECX- co-creation model with the help of appraisal theory. The study suggests that a single interaction of customer with the hotel organization can evoke multiple emotions
Chepngetich *et al.* (2019)	The study broadly identified three factors influencing customer experience in hotels: human clue factor, functional clue factor, mechanic clue factor	All the three factors identified were found to impact customer experience significantly and thereby lead to purchase decision
Ren *et al.* (2016)	Tangible and sensorial experience; staff relational/interactional; aesthetic perception; location	All four factors were significant in predicting the customer experience in budget hotels
Cetin and Walls (2016)	Study explored two major constructs: physical environment and social interaction through a qualitative study of customers and managers	Physical environment and social interactions were further sub-divided with eight factors and these factors were presented to be having a major influence on perceptions and misperceptions pertaining to customer experience
Ali *et al.* (2014)	Four E's of CX: education experience; entertainment experience; aesthetics experience; escapism experience	Customer experience in this study has been defined as a combination of education, entertainment, aesthetics and escapism. All the four E's were found to have positive impact on customer memories and loyalty
Peng *et al.* (2015)	Core attributes, auxiliary attributes, customer experience, revisit intentions	The study confirms both core and auxiliary features were linked to revisit intention; however, only auxiliary attributes were linked to customer experience
Juttner *et al.* (2013)	Cognitive customer responses, emotional customer responses; company-related stimuli	The study used sequential incident laddering technique and found that The proposed method integrates customer and company-related constructs

Note: AR = augmented reality; VR = virtual reality

In terms of immersion, netnography outperforms ethnography and focus groups when it comes to gaining a thorough comprehension of the research topic. Potential customers can use online reviews to determine whether an establishment can be trusted, as well as reduce the risk and ambiguity (Bradley *et al.*, 2016). The hospitality sector is undergoing upheaval because of the advancement of information technology, for example, gamification, digital content and online travel purchases (Sigala, 2015). The netnography approach is quite useful with respect to analyzing tourist's experiences post-tour (Zhang and Hitchcock, 2017). Therefore, netnography was used to study online reviews of branded hotels with the aim of exploring customer experiences of guests during the COVID-19 pandemic.

Selection of observation target group

User-generated content data from TripAdvisor, MakeMyTrip and Booking.com was used as the main corpus for the study. TripAdvisor was selected as a prominent global online information exchange platform (Filieri *et al.*, 2015). It is an online travel company that provides customer feedback on travelers' experiences in hotels, holidays and price comparisons. The second website used for data collection was MakeMyTrip, which is India's largest online travel aggregator. Booking.com is another online travel agency that is widely used in India. It offers hotel bookings and a plethora of unbiased reviews. Few hotels had more reviews on one platform and fewer reviews on another, which balanced the reviews, as customers exhibit preferences for certain platforms for booking or giving feedback. Hotel reviews were collected from these three websites to avoid bias for a particular chain of hotels on an online platform. Figure 1 exhibit the research methodology flow diagram.

Data collection

To represent the best budget hotels, this study chose 12 budget hotels across Delhi NCR, Mumbai, Bengaluru, Kolkata and Chennai, which are the metropolitan cities of India. The screening criteria for the search was set as "Budget Hotels." The search included all budget hotels in metro cities where travelers were more likely to frequent in comparison to non-metro cities. Reviews were collected between April 2020 and April 2021 to capture customer experiences during the COVID-19 pandemic. In India, the COVID-19 lockdown announcement came on March 23, 2020, which led to an environment of chaos and fear. In the initial few months, travel was restricted, and few experiences of quarantined guests were recorded. The objective of this study was to determine the feelings and expectations of customers toward budget hotels and changes in service delivery from budget hotels during the pandemic.

Initially, 2,039 reviews were collected for the 12 hotels; 648 reviews were discarded based on the length and quality of the review, as these reviews were one-line statements or one-word. Non-explanatory reviews, such as "Excellent" and "Our stay was not good," were discarded. Table 2 represents the set of total data collected and the final sample for the study and determinants, including length of the review and reviewer's expertise, are important for the content analysis of customer reviews. In tourism and hospitality services, highly technical knowledge is not required to evaluate the service experience; customers usually compose easy to comprehend reviews without any technical jargon (Huang *et al.*, 2015). Furthermore, only reviews written in English were selected.

Data processing

Thematic analysis was carried out on the collected reviews. According to Braun and Clarke (2006), thematic analysis is a method for identifying, analyzing and reporting patterns (themes) within the data. It is a useful qualitative method to outline the main features of the data (Nowell *et al.*, 2017). Today, researchers prefer Computer-Aided Qualitative Data

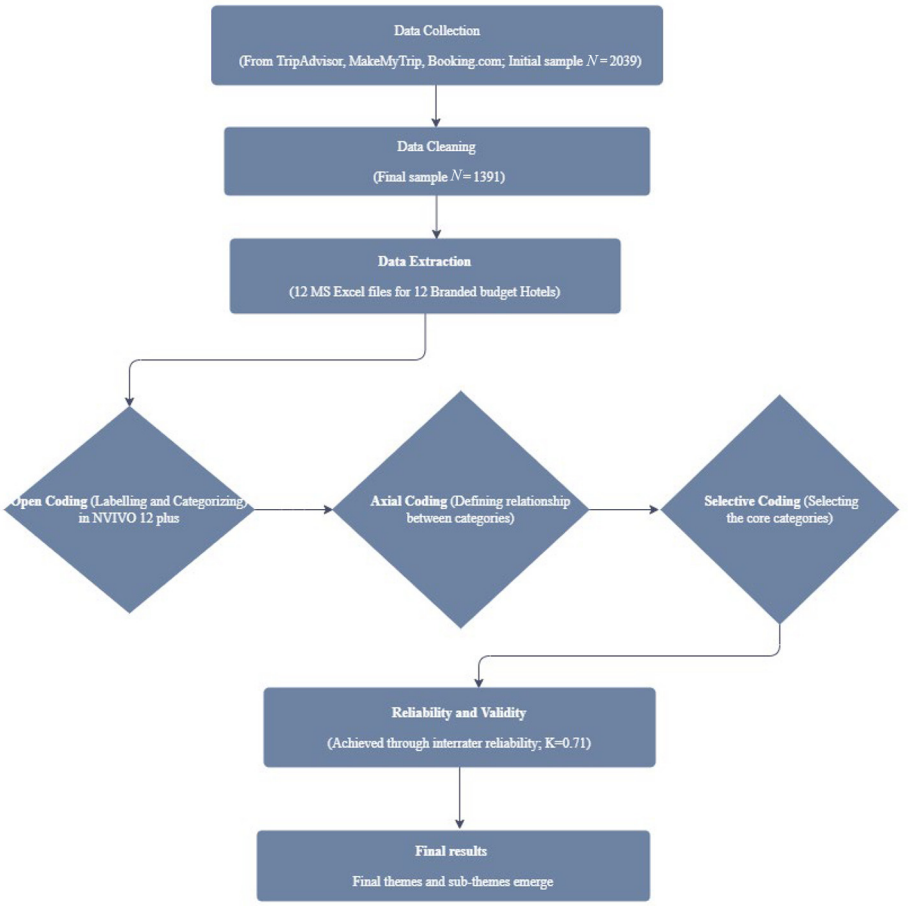


Figure 1.
Research methodology flow diagram

S. no.	Budget hotels	Reviews	Discarded	Final no. of reviews for analysis
1	Lemon Tree Delhi Airport	350	110	240
2	IBIS Mumbai	251	60	191
3	IBIS Kolkata	50	10	40
4	Ginger Mumbai Andheri	49	18	31
5	IBIS Gurugram	107	29	78
6	Sarovar Grand Hometel Mumbai	256	119	137
7	Ginger Chennai	78	19	59
8	The Park Chennai	133	31	102
9	Ginger Bengaluru Whitefield	182	51	131
10	Holiday Inn Banjara hills Hyderabad	242	75	167
11	Holiday Inn New Delhi	71	11	60
12	Holiday Inn Gurugram	270	115	155
		2,039	648	1,391

Table 2.
Sample for the study

Analysis to conduct thematic analysis, as these techniques are accurate, less time consuming and can handle large volumes of data. NVivo software enables the researcher to acquire a greater and more accurate understanding of qualitative data. For data analysis, customer reviews of 12 budget hotels in India were arranged into 12 separate MSWord files. These files were further uploaded to NVivo software. Each of these budget hotels served as "cases," which are "units of analysis." Each case possessed different attributes and attribute values (Jackson and Bazeley, 2019). NVivo 12 Plus was used to analyze the customer reviews. NVivo analyzes qualitative data with the help of nodes, themes and subthemes. Twelve MS-Word files were created for 12 different budget hotel reviews. As the current study was netnographic in nature, there was a limitation on items that could be used as variables. In this study, "trip type" consisted of family, couple, solo and business, and it was not possible to obtain attributes such as travel frequency, education level, income level, age and profession from the travel websites.

After defining the cases and attributes, a coding process was initiated and divided into two steps. The first step was open coding. As this research focused on the customer experience of budget hotels, the customer reviews were uploaded to NVivo 12 Plus to create nodes and codes. A node is a term that serves as a container for grouping words. These terms refer to the same issue or research topic. Different materials containing the same notion are grouped into a node by coding. Transcripts were repeated to decode each different tale and to reconfirm common elements, following the analytic procedure outlined by Ayres *et al.* (2003). The researcher then looked for key statements made by participants that were relevant to the themes. A consensus was reached after several discussions wherein the comments were divided into 10 categories: physical environment, perceived value, hedonic factors, service excellence, personalization, location, food and beverage management, hygiene, customer–customer interaction and employee interaction. The second step involved axial coding. The coded content of the main themes was separately assessed for further analysis. The codes were read repeatedly by all three authors to create themes and subthemes. A cross-case analysis was run to visualize axial codes. Cases occupied the rows, and subcodes occupied the columns. After comprehensive reading, contrasting and re-evaluating, the subthemes were created.

Validity and reliability. Internal reliability is attained when more than one member of a "research team" agrees on what they see and hear (Bryman, 2016, p. 390). The researchers agreed on the data coding and conclusions. The researchers cross-checked 10% of each other's work after coding. Disagreements in the interpretation of reviews were discussed and debated until all three researchers agreed on over 90% of the coding. A debriefing with two peer researchers was done to compare and explain the interpretation of results and classify the primary themes displayed within the transcribed texts, to increase the study's reliability. Finally, key themes emerged from the data analysis, and pertinent quotes were chosen to highlight the underlying narrative. A coefficient of agreement was used to ensure the reliability of the content analysis. This is a standard index with three criteria. The coding process was carried out by two independent coders who worked separately while adhering to the agreed-upon standards. Cohen's Kappa coefficient (K) was calculated, yielding a result of 0.72 and a 97.23% inter-coder agreement, indicating that the themes and essential components were consistent.

Results

Physical environment

Nguyen and Leblanc (2002) defined the physical environment as the surroundings in which the service delivery takes place. The overall decor, spatial layout, artifacts and ambience

form part of a hotel's physical environment (Han and Ryu, 2009). According to a study conducted by Ruiz-Molina *et al.* (2018), the physical environment of a service provider forms an integral part of consumers' evaluation and their satisfaction. During the pandemic, it is natural that consumers would have a certain level of expectation about the room and other related facilities. The way customers perceive hotel stays has changed during the pandemic, and it is seen as quality time spent with family following all safety protocols, rather than as an expensive vacation. The following comments by some guests of a budget hotel reiterate this point:

> Beautifully done suites comfortable cosy room and a breezy balcony. I stayed with my family for two days and I highly recommend this property.

> Lift area was spacious and clean. The lift was sanitised at regular intervals. Toothpicks and tissues were kept operating the lift to avoid direct contact.

Customers interact with the physical ambience more in a hotel setting than in any other industry (Jani and Han, 2014). Hence, the physical environment has a critical role in differentiating service firms from one another and influencing their brand image and end-user behavior (Chen and Hu, 2010; Hu *et al.*, 2009; Ryu *et al.*, 2012). The following review showed how a poor physical environment can lead to customer dissatisfaction:

> [XYZ[1]] hotel needs to work on renovating the rooms. Rooms are very small. Beds are even worse. Mattresses were too soft and couldn't give good night sleep. TVs are too small and was wall mounted at a height which was not properly visible from the bed.

Employee interaction

A major theme emerging among customers of the budget hospitality sector during the pandemic period is the importance of "employee interaction." In a highly competitive hotel industry, the role of employees is considered a key element of business differentiation because it influences customer satisfaction and loyalty toward the service firm. According to Wu (2007), service can be seen as a summation of interactions and performance that happen between customers and employees in the service delivery process. Service employees play a vital role in improving the customer experience, as service cannot be complete without their active participation (Zeithaml *et al.*, 1985). Dimensions such as employee politeness (Parasuraman *et al.*, 1988), employee attentiveness and perceptiveness (Chatzigeorgiou and Simeli, 2017; Farrell *et al.*, 2001; De Ruyter and Wetzels, 2000), listening behavior of employees (Christou and Eaton, 2000; Chandon *et al.*, 1997) and civility and concern (Winsted, 2000) are key to customer satisfaction. This was also evident in our findings:

> Amazing front office team. I am diabetic and had a low sugar episode one night. 'XYZ', I think that was his name, a tall guy, literally saved a trip to the hospital with his quick action.

> The hospitality we received was never seen in any of the hotel it's like something very unique quality and the staffs were excellent. The helping attitude, pleasing words, good body language and gestures were beyond expected. I felt like home here. Undoubtedly, amazing experience.

In contrast, service personnel may be the reason for customer dissatisfaction and complaints if they are not equipped with the right skills, do not supply correct and complete information, do not respond properly or on time or behave in an unacceptable manner (Chan *et al.*, 2017a, 2017b).One guest described a bad experience, stating: "I was travelling with my wife who is a heart patient, and we are both senior citizens. There was nobody to help us

with our luggage despite requesting the reception and it was so difficult to get from parking to rooms. Very bad experience."

However, the COVID-19 pandemic has put restrictions on personal interactions (Chan *et al.*, 2020; WHO, 2020a); this has accelerated the use of unmanned devices, limiting human-to-human contact.

Hedonic factors

Hedonic factors can be understood as the activities and experiences that relate to one's sensory pleasures, aesthetics, enjoyment and fantasies as part of the consumption process (Holbrook and Hirschman, 1982; Babin *et al.*, 1994; Kempf, 1999).In this study, customers' decisions regarding their stay were closely linked to their desire for relaxation after the first phase of the lockdown. Additionally, the analysis of customer reviews revealed their focus on hedonic factors that they experienced during their stay. One customer commented:

> Amenities are good and the vibe is cool. Cricket, Badminton and Gym facilities are superb to add on to your stay. The hotel had all required facilities and I enjoyed using the badminton court in the morning.

Hedonic services are consumed to achieve affective gratification (Kempf, 1999) and are of specific relevance to the hospitality sector due to consumers' involvement in the overall purchase and consumption process (Gilmore and Pine, 2002; Mowen, 1987; Titz, 2009). During the pandemic, managers were challenged to introduce, renovate and refurbish their core service values to stay relevant and follow all COVID-19 protocols. However, the costs associated with COVID-19 prevention measures and upgrading work processes, such as the implementation of new operating protocols and adoption of digital tools, may have also been relatively higher for budget hotels. The following comment by a customer demonstrated the efforts taken by a budget hotel to ensure the safety of guests while offering pleasure and relaxation:

> Travel and roaming around the city were arranged by travel desk of the hotel following all protocols laid down by the Government. They ensured that the guest followed all precautions like Use of Face Mask, Hand Sanitisers and separate vehicles for each family. They provided a kit which contained masks, sanitisers, wet wipes, and water bottles.

Of the emotions displayed by people during the pandemic, two types of emotions were found to be very prominent: extreme fear and uncertainty (Esterwood and Saeed, 2020). Evidence gathered post-lockdown considers some factors, namely the fear of going out, fear of infection "post-lockdown," anxiety, fear of infecting others and fear of returning to the new normal (Goolsbee and Syverson, 2021). However, Larson and Shin (2018) suggested that shoppers who were under the greatest emotional stress were also the ones who were more likely to engage in hedonic and utilitarian shopping post-pandemic. The imposed social distancing had forced people to stay at home during the lockdown period. Generally, people love to socialize and were looking forward to initiating travel (Mohd Dali *et al.*, 2020) after the first wave of restrictions.

With the COVID-19 pandemic continuing to impact occupancy rates of budget hotels, hoteliers are doing whatever they can to communicate with travelers that they are taking every precaution to make their stay safe and offer the best options for enjoyment. The narratives observed in this study showed that travelers were unsure of which precautions were impactful. The following opinion of a customer narrates it.

> Our kids love the outdoor games, although it was open on our last visit, we restricted our movement within the room and restaurant only due to COVID 19.

Location

The location factor is more specific to the hotel industry than any other service industry. As the production and consumption of services happen simultaneously in the service industry, the success of the hotel industry relies heavily on location (Nozar, 2001; Pan, 2005). Xiang and Krawczyk (2016) noted that the location of the hotel has a major impact on guest satisfaction. One satisfied guest specifically commented about the location of the hotel:

> I prefer staying here whenever I visit [XYZ] and the property stands out because of its convenient location and proximity to the places I frequently visit during my stay.

The existing literature suggests that based on the purpose of the customer's travel, their emphasis on the location of the hotel differs (Chan and Wong, 2006; Walker, 2008; Yavas and Babakus, 2005). This is consistent with the findings of our study. Customers who travelled during the pandemic preferred to stay in hotels that were close to the airport/metro stations and other facilities. A few comments from the reviewers demonstrates this trend:

> This hotel is very close to the airport and takes 5-10 minutes by a cab, and I think that's the biggest locational advantage they have. There are shops, restaurants and pharmacy at a 5–10-minute walkable distance so it's not that difficult to go and get anything that you need.

> The location is easy to reach, its right on the main road near the metro station. Very convenient location.

Self-service technologies

Big budget hotel chains are now using self service technologies (SSTs) in the form of kiosks, mobile phones, tablets, laptops and installing technology that enables customers to self-check-in/out and book tables, among other things. Self-service technologies (SST) enable customers to produce and deliver service independent of direct service employee involvement (Meuter *et al.*, 2000) and are an alternative or a replacement for human-based service transactions (Rust and Espinoza, 2006). One customer commented:

> I could order food online from the restaurant by scanning quick response code, which was available in my room, and I could get a real-time visibility into the menu, along with the dishes that are available and out of stock. Amazing facility.

The use of technology was inevitable during the COVID-19 pandemic to maintain social distancing norms and contain the spread of the virus (Peña Ruiz, 2020). The hotel staff can perform a number of tasks in a timely and efficient manner using devices connected to the internet, including retrieving bookings, taking payments online, facilitating the ability of guests to choose their room from a digital interface, availability for virtual chatting and providing virtual tours for guests to learn about the hotel and surrounding area (Cheong *et al.*, 2017). Service managers sought technology interfaces to maintain service quality and enhance experiences while initiating the COVID-19 protocols proposed by health authorities (Nguyen and Vu, 2020; Sigala, 2020). SSTs were found to be useful in avoiding physical contacts and also they help save time. Customers can use these technologies without anyone's assistance; however, these technologies may be a hindrance for people who are not comfortable using them or when the technology does not resolve their concern (Sigala, 2020). There were many comments about the convenience of the facilities. A guest of one hotel recounted:

> Contactless check-in/checkout facility provided by this hotel was smooth. We could save the time and also feel safe during the COVID time.

While efforts were made to offer a safe experience by using technology, some customers experienced problems during their stay. One customer noted:

> I do not recommend [XYZ]hotel. We were made to wait for 30 minutes to check-in since their online system was not down. To make it worse they made me to pay by cash.

Personalization

Personalization as an act allows customers to be presented with information based on their own preferences or profiles. As it is believed that every user is unique, some people feel that their information preferences are unique as well (Dangi and Malik, 2017). Personalization also means the continuous improvement of guest experiences by satisfying their needs effectively and efficiently (Ferenczuk, 2018). The hospitableness of hotel services can best be explained by personalization dimensions, as suggested by Ariffin and Maghzi (2012). Hemmington (2007) opined that hospitality should aim to provide guests with experiences that are not only unforgettable and enriching but also personal in nature. Personalization is considered a key factor in premium hotels, although it is not considered an important feature of budget hotels. However, personal care and attention were extremely crucial for guests who visited hotels during the COVID-19 period, regardless of their status as a premium or budget hotel. This was demonstrated in the following comments:

> I was quarantined at this hotel. The care provided to us by the hotel was great. The staff was lovely and accommodating, we were very impressed and felt very well looked after.

> The hotel provided Sanitizes and Masks along with other goodies in the welcome basket. This was a nice gesture.

Service personalization, which is carried out by explaining the effect of individual context on customer needs and accordingly adjusting the service attributes, can co-create high value and satisfaction (Choi *et al.*, 2019; Massimo and Ricci, 2019).

According to Ariffin and Maghzi (2012), personalization refers to the guests¨ feelings about how successful the hotel was in treating them as important guests and is important at every interaction in the customer's journey because even a small thing can influence the customer's experience (Bizcommunity.com, 2018). The following comment showcases a customer's dissatisfaction:

> To clean my room I used to call housekeeping staff continuously from the morning, at last they cleaned my room in the evening. I tried calling help desk but No one was there to take care of my requirements.

Hygiene

The COVID-19 pandemic has demonstrated that hygiene factors are more important now than ever to consumers (Gossling *et al.*, 2020). Sanitation and hygiene factors weigh very high in the current situation, regardless of age (Kock *et al.*, 2020). Recent research suggested that in the ongoing pandemic, hygiene and sanitation factors influence consumer's decisions pertaining to accommodations (Gossling *et al.*, 2020). Hygiene factors also include safety compliance measures as directed by health-care and workplace safety authorities (Kubickova and Neal, 2000). All hotels have adopted their own standard operating procedures to maintain high standards of hygiene. A few comments from our study also clearly indicated the processes followed by budget hotels in India.

> As soon as I reached the hotel and parked my car, the security staff did thermal screening, noted our details and sanitized our bags. There is a good arrangement for sanitizing the footwear too at the entrance. There is also proper glass partition at the reception where you have to stand and get the room allotted.

The complimentary breakfast provided here is in a disposable plate, which is safe in this condition of COVID.

The notion of hygiene encompasses much more than cleanliness and sanitation and includes broader aspects of health care. During the pandemic, customers prepare a travel plan to ensure their health and safety while they are working, which includes specific details like seeking medical help if they fall ill (Jiang and Wen, 2020). The following comment by a customer of a budget hotel showed the importance of health-care-related aspects during their stay:

> The staff provided the details of hospitals nearby and also the COVID 19 testing centers in case of any requirement.

Perceived value

Perceived value is considered to be an essential contributing factor to the success of a business and an important measure of customer loyalty (Kim *et al.*, 2015).Perceived value is defined in monetary terms and has been shown to have four dimensions:

(1) low price for the service;

(2) presence of whatever one wants in a product;

(3) quality that the consumer receives for the price paid; and

(4) what the consumer gets for what they give in the form of price (Zeithaml, 1988).

In the hospitality sector, the value and quality of the service, as perceived by the customer, are vital (Joung *et al.*, 2016; Sabiote-Ortiz *et al.*, 2016; Zeleti *et al.*, 2016). Although the rapid spread of infection led to the loss of lives and jobs, the immediate financial impact of COVID-19 impacted the way people perceive pleasure and enjoyment, making them more financially cautious. One customer commented:

> Considering the average amenities offered by [XYZ] hotel, I think my decision to stay in this hotel was not good. The rent was too high and will not recommend this hotel for a family.

Discussion

Cluster analysis was performed to cross-check and validate the final themes. The themes were then used to create a hierarchical chart (Figure 3). The hierarchical chart indicates personalization played the most important role in contributing to the customer experience, followed by hygiene. Personalization, as observed from the customer reviews, indicated special requests with respect to COVID-19 protocols, food and beverages/housekeeping services and late checkouts. In addition, some guests quarantined in these budget hotels, and for them, personalization or customization was the key element, as quarantining in a single room for more than 12 days takes a toll on an individual's emotional and psychological well-being. Many guests mentioned how extra attention was given to quarantining guests, as they demanded personalized services, whether it was with respect to food and beverage or housekeeping or medical help. In addition, the staff was found to have gone the extra mile to deliver satisfaction to these guests. The current finding is in sync with the previous studies (Ariffin *et al.*, 2018, Ren *et al.*, 2016; Mcintosh and Siggs, 2005) where

personalization was found to be the predictor of guest satisfaction; however, these studies were carried in different hotel settings like luxury, boutique accommodation.

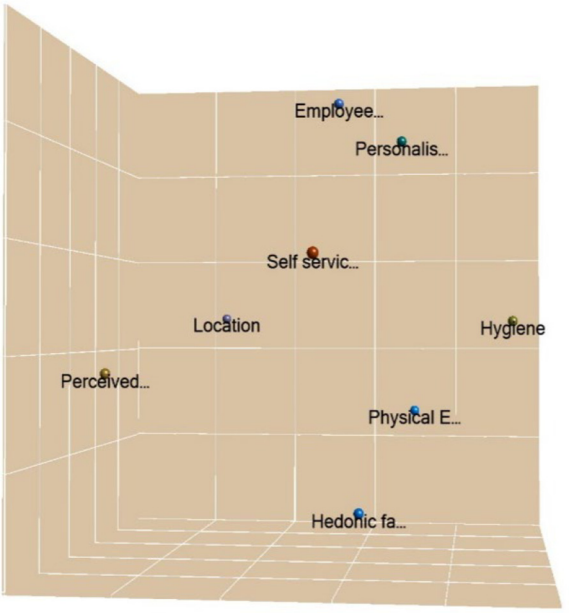
In the past, some studies have compared hedonic and utilitarian motivations among hotel guests (Nguyen and Vu, 2020). Hedonic experience as an intangible benefit is an essential component of luxury consumption (Zhang *et al.*, 2020; Kim *et al.*, 2020), and further, it is expected that the hedonic value perceptions of hotel guests would positively affect their behavioral intentions. (Dedeoğlu and Boğan, 2021). As per the literature, hedonic factors were mentioned less frequently by the hotel guests of budget hotels. In general, budget hotels have less hedonic facilities, and they tend to be more utilitarian, aiming to provide good food, decent ambience and great customer service but the finding of our study i.e. the growing demand for hedonic experiences even while staying in budget properties is a reflection of growing aspirations of the budget segment customer. The hierarchical chart (Figure 3) also reflected that although SSTs emerged as a major theme, many customers expected more in terms of touchless entry and payments, scanning options for food menu and other facilities. Budget hotels must empower their customers to interact digitally, whether through digital kiosks, mobile apps or biometric data, in terms of how they engage. This outcome is supported by a study conducted by Ren *et al.* (2016) suggesting adoption of service technology in budget hotels. Keeping guests' information private and secure, yet offering personalized services with the latest technological interventions, is necessary. To preserve social distancing, many hotel chains offer check-in via mobile phones, which act as "virtual queue systems" where personalized requests can also be logged. Although budget hotels cannot offer much within the stipulated budget, the segment surely needs to work upon creating a touchless customer experience using the latest technologies (Figure 2 and 3).

Figure 4 reflects the preferences of different travelers, demonstrating that hygiene and personalization were the key elements of the customer experience across budget hotels. For solo travelers, the perceived value mattered the most, as they typically travel within a stipulated budget, while in the "friends" category, personalization was the most important theme. Couple travelers valued the physical environment more than any other traveler group. In previous studies, the hygiene factor was one of the factors but was not the most important factor. Staff training and hygiene checks hold extreme significance with respect to "new normalcy" travel prerequisites. Hence, management needs to prioritize hygiene factors and protection measures. Investing in tangible and intangible elements would help elevate the holistic physical environment and social environment. Customer journey elements starting from the booking page website, providing necessary information and building trust on digital touchpoints are also quite significant. Signs like "COVID protected" and "COVID safe" on hotel websites and Online travel agency websites can build trust and reduce the perceived risk, thus increasing the likelihood of a customer making a purchase. While previous studies conducted in the area of budget hotels highlight return-on-Investment (ROI) being the most important factor, the current study brings personalization and hygiene as the supreme actors affecting customer experience.

Implications

Theoretical implications

This study contributes to the literature in the field of customer experience research in the hospitality sector by extending previous research in the backdrop of COVID-19 pandemic (Filimonau *et al.*, 2020; Seyitoglu and Ivanov, 2020; Jiang and Wen, 2020; Shin and Kang, 2020). First, a major theoretical contribution of this study was the use of netnography to understand dimensions of the customer experience. Though the netnography approach is



Source: Created by author in NVivo 12 plus

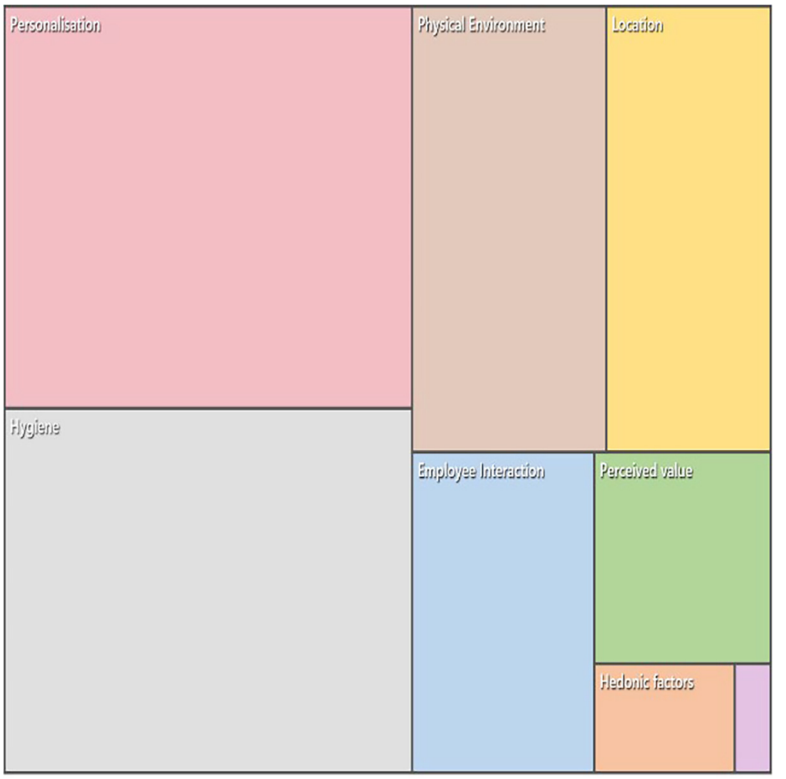
Figure 2.
Cluster analysis

slowly becoming one of the popular qualitative research techniques in marketing only few researchers have adopted this method as a tool for qualitative study (Brown *et al.*, 2015; Kozinets and Handelman, 1998). The insight into the customers views after using the services of budget hotels during the pandemic was emphasized through our use of the unobtrusive data collection method. As our study used reviews posted by actual guests who used the services of budget hotels, we argue that the findings present higher levels of reliability.

Second outcome of this study is the identification of key factors contributing to the overall customer experience of guests staying in budget hotels during COVID-19 pandemic. The netnography approach facilitated us to use freely available information sources to derive meaningful outcomes about the customer experience by examining and categorizing the views of customers about their stay and analyzing the impact of these measures on their intended customer experience. Though factors that emerged from our analysis are consistent with the existing literature (Ruiz-Molina *et al.*, 2018; Xiang and Krawczyk, 2016), the priorities assigned by the guests to these factors during a health crisis such as COVID 19 is very different. Some of the factors seem to have become more important like personalization, hygiene, SSTs, employee interaction considering COVID-19 protocols. It is important to note that while personalization and SSTs were highlighted much in the context of luxury hotels prior to pandemic (Ariffin and Maghzi, 2012; Haruzivishe, 2014; Wilkins *et al.*, 2007), it was never a top priority for guests staying in budget hotels.

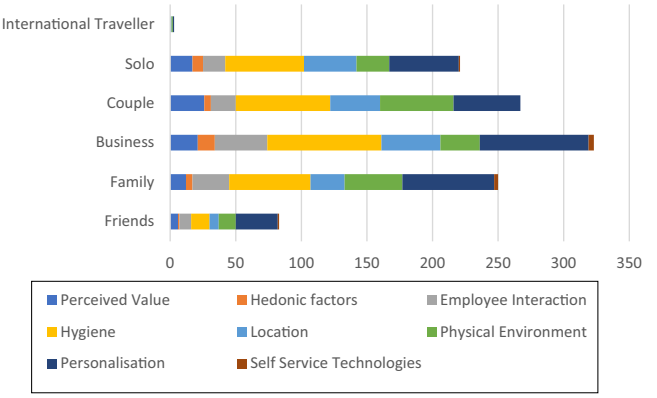
Managerial implications

Pandemics like COVID-19 are uninvited and create an unsettling environment for businesses, necessitating the expansion of all available resources such as land, technology, money, infrastructure and employees (Thomas *et al.*, 2022). The budget hotels need to



Source: Created by author in NVivo 12 plus

Figure 3.
Hierarchical chart (thematic analysis)



Source: Created by author in MS-Excel

Figure 4.
Theme significance as per traveler segment

showcase agility in terms of preparation, risk assessment and implementing quick infrastructure. Therefore, the study offers some key insights for managers of budget hotels and policymakers. Personalization was previously not an important factor for a guest staying in a budget hotel (Cetin and Walls, 2016). However, due to the current COVID-19 restrictions and fear of infection, service personalization emerged as the most important attribute for the customer experience. Managers should consider using the data provided by the guests for personalized offerings, as their counterparts have done in luxury hotels. Even though the pandemic has put restrictions on frequent employee–customer interactions, the manager and employees should extend personalization in the overall process of service delivery and should consider each interaction an opportunity to create memories.

An important aspect for the managers to consider is enhancing the customer experience by using SSTs like facial scanners, touchless contact devices, Web check-in to enable the guests to go through contactless processes. In addition to this our findings revealed that consumers these days pay more attention to health and wellness – both physical health and mental health. Budget hotels can offer customized healthy diets, technologically savvy equipment to measure their health parameters and tie up with doctors, psychologists to take care of the guests' physical/mental wellness specially in times of crisis. These factors are important to gain the trust of the guests about the various initiatives introduced by budget hotels during the pandemic, culminating in a safe stay. A clear understanding of the dimensions of the customer experience of budget hotels in India is important for its sustainability and longevity because budget hotels is an important segment in India.

The role of online reviews in the hospitality sector is vital because it enables tourists to share experiences and opinions (AmArAl *et al.*, 2014, Lu and Stepchenkova, 2015). This also forms a potential source of reference for customers in their future decisions-making process. This study recommends frequent analysis of online reviews by budget hotels to redesign customer experience considering the extensive use of online portals and other travel related platforms. The need for a complete redesign was very evident in our analysis.

Limitations and future research directions

The main limitation of this study relates to the low generalizability of the results because this was conducted in the context of India and during the ongoing pandemic. Further, the social and demographic characteristics of the guests who posted the reviews were not considered. Additionally, netnography only considers customer reviews that are posted online. Thus, this analysis ignored the views of customers who did not post their comments anywhere. The results might change considering other review types. Therefore, care should be given when interpreting these results.

Given the importance of redesigning the customer experience in budget hotels due to pandemic regulations, it is important to understand the views of customers and managers. The current study is an attempt to understand the views of customers and ignores managers' perspectives. It would be interesting for future researchers to examine all aspects of the customer experience from the managers' point of view. Further research can be done on a larger sample with other countries' budget hotels to validate the extracted dimensions. Another avenue for future research is to conduct quantitative study on the proposed dimensions to measure customer experience of budget hotels. Also, more such studies on customer experience can be conducted in other service sectors against the backdrop of the COVID-19 pandemic, which will help organizations to be more resilient with pragmatic strategies in situation of crisis.

References

Adjei, M., Noble, S. and Noble, C. (2009), "The influence of C2C communications in online brand communities on customer purchase behavior", *Journal of the Academy of Marketing Science*, Vol. 38 No. 5, pp. 634-653, doi: 10.1007/s11747-009-0178-5.

Ali, F., Hussain, K. and Ari Ragavan, N. (2014), "Memorable customer experience: examining the effects of customers experience on memories and loyalty in Malaysian resort hotels", *Procedia – Social and Behavioral Sciences*, Vol. 144, pp. 273-279.

AmArAl, F., TiAgo, T. and TiAgo, F. (2014), "User-generated content: tourists' profiles on Tripadvisor", *International Journal of Strategic Innovative Marketing*, Vol. 1 No. 3, pp. 137-145.

Anguera-Torrell, O., Aznar-Alarcón, J.P. and Vives-Perez, J. (2021), "COVID-19: hotel industry response to the pandemic evolution and to the public sector economic measures", *Tourism Recreation Research*, Vol. 46 No. 2, pp. 148-157.

Ariffin, A.A.M. and Maghzi, A. (2012), "A preliminary study on customer expectations of hotel hospitality: influences of personal and hotel factors", *International Journal of Hospitality Management*, Vol. 31 No. 1, pp. 191-198.

Ariffin, S.K., Mohan, T. and Goh, Y.N. (2018), "Influence of consumers' perceived risk on consumers' online purchase intention", *Journal of Research in Interactive Marketing*.

Ayres, L., Kavanaugh, K. and Knafl, K.A. (2003), "Within-case and across-case approaches to qualitative data analysis", *Qualitative Health Research*, Vol. 13 No. 6, pp. 871-883, doi: 10.1177/1049732303013006008.

Babin, B.J., Darden, W.R. and Griffin, M. (1994), "Work and/or fun: measuring hedonic and utilitarian shopping value", *Journal of Consumer Research*, Vol. 20 No. 4, pp. 644-656.

Bartik, A.W., Bertrand, M., Cullen, Z.B., Glaeser, E.L., Luca, M. and Stanton, C.T. (2020), "How are small businesses adjusting to COVID-19? Early evidence from a survey (No. w26989)", National Bureau of Economic Research.

Bharwani, S. and Mathews, D. (2021), "Post-pandemic pressures to pivot: tech transformations in luxury hotels", *Worldwide Hospitality and Tourism Themes*, Vol. 13 No. 5.

Bonfanti, A., Vigolo, V. and Yfantidou, G. (2021), "The impact of the Covid-19 pandemic on customer experience design: the hotel managers' perspective", *International Journal of Hospitality Management*, Vol. 94, p. 102871.

Bradley, G.L., Sparks, B.A. and Weber, K. (2016), "Perceived prevalence and personal impact of negative online reviews", *Journal of Service Management*, Vol. 27 No. 4, pp. 507-533, doi: 10.1108/JOSM-07-2015-0202.

Braun, V. and Clarke, V. (2006), "Using thematic analysis in psychology", *Qualitative Research in Psychology*, Vol. 3 No. 2, pp. 77-101.

Brotherton, B. (2004), "Critical success factors in UK budget hotel operations", *International Journal of Operations and Production Management*, Vol. 24 No. 9.

Brown, E.A., Thomas, N.J. and Bosselman, R.H. (2015), "Are they leaving or staying: a qualitative analysis of turnover issues for Generation Y hospitality employees with a hospitality education", *International Journal of Hospitality Management*, Vol. 46, pp. 130-137.

Brunner-Sperdin, A., Peters, M. and Strobl, A. (2012), "It is all about the emotional state: managing tourists' experiences", *International Journal of Hospitality Management*, Vol. 31 No. 1, pp. 23-30.

Bryman, A. (2016), *Social Research Methods*, Oxford University press, New York, NY.

Cetin, G. and Walls, A. (2016), "Understanding the customer experiences from the perspective of guests and hotel managers: empirical findings from luxury hotels in Istanbul", *Journal of Hospitality Marketing and Management*, Vol. 25 No. 4, pp. 395-424.

Chan, G.S.H., Tang, I.L.F. and Sou, A.H.K. (2017b), "An exploration of consumer complaint behavior towards the hotel industry: case study in Macao", *International Journal of Marketing Studies*, Vol. 9 No. 5, pp. 56-76.

Chan, K.W., Gong, T., Zhang, R. and Zhou, M. (2017a), "Do employee citizenship behaviors lead to customer citizenship behaviors? The roles of dual identification and service climate", *Journal of Service Research*, Vol. 20 No. 3, pp. 259-274.

Chan, I.C.C., Ma, J., Law, R., Buhalis, D. and Hatter, R. (2020), "Dynamics of hotel website browsing activity: the power of informatics and data analytics", *Industrial Management and Data Systems*, Vol. 121 No. 6.

Chandon, J.L., Leo, P.Y. and Philippe, J. (1997), "Service encounter dimensions-a dyadic perspective: measuring the dimensions of service encounters as perceived by customers and personnel", *International Journal of Service Industry Management*, Vol. 8 No. 1.

Chatzigeorgiou, C. and Simeli, I. (2017), "Perception of service quality in agrotourism accommodations: impact on guest loyalty and re-visit intentions", *Journal of Tourism, Heritage and Services Marketing (JTHSM)*, Vol. 3 No. 1, pp. 33-41.

Chen, S.H., Tzeng, S.Y., Tham, A. and Chu, P.X. (2021), "Hospitality services in the post COVID-19 era: are we ready for high-tech and no touch service delivery in smart hotels?", *Journal of Hospitality Marketing and Management*, Vol. 30 No. 8, pp. 905-928.

Cheong, S.N., Ling, H.C., Teh, P.L., Ahmed, P.K. and Yap, W.J. (2017), "Encrypted quick response scheme for hotel check-in and access control system", *International Journal of Engineering Business Management*, Vol. 9.

Chepngetich, B., Ouma, O.K. and Aila, F.O. (2019), "Influence of customer experience dimensions on purchase behaviour in Kenyan hotels", *European Journal of Business and Management*, Vol. 11 No. 21, doi: 10.7176/EJBM.

Chittiprolu, V., Samala, N. and Bellamkonda, R.S. (2021), "Heritage hotels and customer experience: a text mining analysis of online reviews", *International Journal of Culture, Tourism and Hospitality Research*, Vol. 15 No. 2.

Choi, T.M., Ma, C., Shen, B. and Sun, Q. (2019), "Optimal pricing in mass customization supply chains with risk-averse agents and retail competition", *Omega*, Vol. 88, pp. 150-161.

Christou, E. and Eaton, J. (2000), "Management competencies for graduate trainees", *Annals of Tourism Research*, Vol. 27 No. 4, pp. 1058-1061.

Dangi, H. and Malik, A. (2017), "Personalisation in marketing: an exploratory study", *International Journal of Internet Marketing and Advertising*, Vol. 11 No. 2, pp. 124-136.

Davari, D. and Jang, S. (2021), "Visit intention of non-visitors: a step toward advancing a people-centered image", *Journal of Destination Marketing & Management*, Vol. 22, p. 100662.

Davari, D., Vayghan, S., Jang, S.S. and Erdem, M. (2022), "Hotel experiences during the COVID-19 pandemic: high-touch versus high-tech", *International Journal of Contemporary Hospitality Management*, Vol. 34 No. 4.

De Ruyter, K. and Wetzels, M. (2000), "Customer equity considerations in service recovery: a cross-industry perspective", *International Journal of Service Industry Management*, Vol. 11 No. 1.

Dedeoğlu, B.B. and Boğan, E. (2021), "The motivations of visiting upscale restaurants during the COVID-19 pandemic: the role of risk perception and trust in government", *International Journal of Hospitality Management*, Vol. 95, p. 102905.

Díaz, M.R. and Rodríguez, T.F.E. (2018), "Determining the reliability and validity of online reputation databases for lodging: Booking. com, TripAdvisor, and HolidayCheck", *Journal of Vacation Marketing*, Vol. 24 No. 3, pp. 261-274.

Escandon-Barbosa, D. and Salas-Paramo, J. (2021), "Tourism amidst COVID-19: consumer experience in luxury hotels booked through digital platforms", *Tourism Recreation Research*, pp. 1-6, doi: 10.1080/02508281.2021.1940708.

Esterwood, E. and Saeed, S.A. (2020), "Past epidemics, natural disasters, COVID19, and mental health: learning from history as we deal with the present and prepare for the future", *Psychiatric Quarterly*, Vol. 91 No. 4, pp. 1121-1133.

Farrell, A.M., Souchon, A.L. and Durden, G.R. (2001), "Service encounter conceptualisation: employees' service behaviours and customers' service quality perceptions", *Journal of Marketing Management*, Vol. 17 Nos 5/6, pp. 577-593.

Ferenczuk, K. (2018), "Personalisation in hospitality industry", Hotelier.com., available at: https://insights.ehotelier.com/insights/2018/04/30/personalisation-hospitality-industry/ (accessed 28 October 2021).

Filieri, R., Alguezaui, S. and McLeay, F. (2015), "Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth", *Tourism Management*, Vol. 51, pp. 174-185, doi: 10.1016/j.tourman.2015.05.007.

Filimonau, V., Derqui, B. and Matute, J. (2020), "The COVID-19 pandemic and organisational commitment of senior hotel managers", *International Journal of Hospitality Management*, Vol. 91, p. 102659.

Fiorentino, A. (1995), "Budget hotels: not just minor hospitality products", *Tourism Management*, Vol. 16 No. 6, pp. 455-462.

Gilmore, J.H. and Pine, B.J. (2002), "Differentiating hospitality operations via experiences: why selling services is not enough", *The Cornell Hotel and Restaurant Administration Quarterly*, Vol. 43 No. 3, pp. 87-96.

Godovykh, M. and Tasci, A.D. (2020), "Customer experience in tourism: a review of definitions, components, and measurements", *Tourism Management Perspectives*, Vol. 35, p. 100694.

Goolsbee, A. and Syverson, C. (2021), "Fear, lockdown, and diversion: comparing drivers of pandemic economic decline 2020", *Journal of Public Economics*, Vol. 193, p. 104311.

Gössling, S., Scott, D. and Hall, C.M. (2020), "Pandemics, tourism and global change: a rapid assessment of COVID-19", *Journal of Sustainable Tourism*, Vol. 29 No. 1, pp. 1-20.

Gunasekar, S., Kumar, D.S., Purani, K., Sudhakar, S., Dixit, S.K. and Menon, D. (2021), "Understanding service quality attributes that drive user ratings: a text mining approach", *Journal of Vacation Marketing*, Vol. 27 No. 4, pp. 400-419.

Gupta, S. and Vajic, M. (2000), "The contextual and dialectical nature of experiences", *New Service Development: Creating Memorable Experiences*, Vol. 15, pp. 33-51.

Gyódi, K. (2021), "Airbnb and hotels during COVID-19: different strategies to survive", *International Journal of Culture, Tourism and Hospitality Research*, Vol. 16 No. 1.

Han, H. and Hyun, S.S. (2017), "Impact of hotel-restaurant image and quality of physical-environment, service, and food on satisfaction and intention", *International Journal of Hospitality Management*, Vol. 63, pp. 82-92.

Han, H. and Ryu, K. (2009), "The roles of the physical environment, price perception, and customer satisfaction in determining customer loyalty in the restaurant industry", *Journal of Hospitality & Tourism Research*, Vol. 33 No. 4, pp. 487-510.

Hao, F., Xiao, Q. and Chon, K. (2020), "COVID-19 and China's hotel industry: impacts, a disaster management framework, and post-pandemic agenda", *International Journal of Hospitality Management*, Vol. 90, p. 102636.

Hao, F. and Chon, K. (2021), "Are you ready for a contactless future? A multi-group analysis of experience, delight, customer equity, and trust based on the technology readiness index 2.0", *Journal of Travel and Tourism Marketing*, Vol. 38 No. 9, pp. 900-916.

Haruzivishe, F. (2014), "Use of facial recognition for data personalisation in customer relationship management: case of great Zimbabwe hotel", (Doctoral dissertation, Bindura University of Science Education).

Hemmington, N. (2007), "From service to experience: understanding and defining the hospitality business", *The Service Industries Journal*, Vol. 27 No. 6, pp. 747-755.

Holbrook, M.B. and Hirschman, E.C. (1982), "The experiential aspects of consumption: consumer fantasies, feelings, and fun", *Journal of Consumer Research*, Vol. 9 No. 2, pp. 132-140.

Hu, H. H., Kandampully, J. and Juwaheer, T. D. (2009), "Relationships and impacts of service quality, perceived value, customer satisfaction, and image: an empirical study", *The Service Industries Journal*, Vol. 29 No. 2, pp. 111-125.

Huang, A., Chen, K., Yen, D. and Tran, T. (2015), "A study of factors that contribute to online review helpfulness", *Computers in Human Behavior*, Vol. 48, pp. 17-27, doi: 10.1016/j.chb.2015.01.010.

Jackson, K. and Bazeley, P. (2019), *Qualitative Data Analysis with NVivo*, 3rd ed., SAGE Publications Ltd, London.

Jani, D. and Han, H. (2014), "Personality, satisfaction, image, ambience, and loyalty: testing their relationships in the hotel industry", *International Journal of Hospitality Management*, Vol. 37, pp. 11-20.

Ji, M., Li, M. and Hsu, C.H. (2016), "Emotional encounters of Chinese tourists to Japan", *Journal of Travel and Tourism Marketing*, Vol. 33 No. 5, pp. 645-657, doi: 10.1080/10548408.2016.1167353.

Jiang, Y. and Wen, J. (2020), "Effects of COVID-19 on hotel marketing and management: a perspective article", *International Journal of Contemporary Hospitality Management*, Vol. 32 No. 8.

Joung, H.W., Choi, E.K. and Wang, E. (2016), "Effects of perceived quality and perceived value of campus foodservice on customer satisfaction: moderating role of gender", *Journal of Quality Assurance in Hospitality and Tourism*, Vol. 17 No. 2, pp. 101-113.

Jüttner, U., Schaffner, D., Windler, K. and Maklan, S. (2013), "Customer service experiences: developing and applying a sequential incident laddering technique", *European Journal of Marketing*, Vol. 47 Nos 5/6.

Kaushal, V. and Yadav, R. (2020), "Understanding customer experience of culinary tourism through food tours of Delhi", *International Journal of Tourism Cities*.

Kempf, D.S. (1999), "Attitude formation from product trial: distinct roles of cognition and affect for hedonic and functional products", *Psychology and Marketing*, Vol. 16 No. 1, pp. 35-50.

Kim, J.H. and Ritchie, J.B. (2014), "Cross-cultural validation of a memorable tourism experience scale (MTES)", *Journal of Travel Research*, Vol. 53 No. 3, pp. 323-335.

Kim, J.J., Choe, J.Y.J. and Hwang, J. (2020), "Application of consumer innovativeness to the context of robotic restaurants", *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 1.

Kim, Y.H., Duncan, J. and Chung, B.W. (2015), "Involvement, satisfaction, perceived value, and revisit intention: a case study of a food festival", *Journal of Culinary Science and Technology*, Vol. 13 No. 2, pp. 133-158.

Klaus, P. and Maklan, S. (2007), "The role of brands in a service-dominated world", *Journal of Brand Management*, Vol. 15 No. 2, pp. 115-122.

Kock, F., Nørfelt, A., Josiassen, A., Assaf, A.G. and Tsionas, M.G. (2020), "Understanding the COVID-19 tourist psyche: the evolutionary tourism paradigm", *Annals of Tourism Research*, Vol. 85, p. 103053.

Kozinets, R. (2015), "Management netnography: axiological and methodological developments in online cultural business research", *The Sage Handbook of Qualitative Business and Management Research Methods*, Sage, London.

Kozinets, R.V. and Handelman, J. (1998), "Ensouling consumption: a netnographic exploration of the meaning of boycotting behavior, *ACR North American Advances*, Vol. 25, pp. 475-480.

Kubickova, M. and Neal, R. (2020), "Challenges and opportunities with H-2B visas: can luxury resorts survive without them?", *International Journal of Contemporary Hospitality Management*, Vol. 33 No. 1.

Kuo, Y.F., Wu, C.M. and Deng, W.J. (2009), "The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services", *Computers in Human Behavior*, Vol. 25 No. 4, pp. 887-896.

Larsen, S. (2007), "Aspects of a psychology of the tourist experience", *Scandinavian Journal of Hospitality and Tourism*, Vol. 7 No. 1, pp. 7-18.

Larson, L.R. and Shin, H. (2018), "Fear during natural disaster: its impact on perceptions of shopping convenience and shopping behavior", *Services Marketing Quarterly*, Vol. 39 No. 4, pp. 293-309.

Lee, K. M., Park, N. and Song, H. (2005), "Can a robot be perceived as a developing creature? Effects of a robot's long-term cognitive developments on its social presence and people's social responses toward it", *Human Communication Research*, Vol. 31 No. 4, pp. 538-563.

Lemon, K.N. and Verhoef, P.C. (2016), "Understanding customer experience throughout the customer journey", *Journal of Marketing*, Vol. 80 No. 6, pp. 69-96.

Li, M., Yin, D., Qiu, H. and Bai, B. (2021), "A systematic review of AI technology-based service encounters: implications for hospitality and tourism operations", *International Journal of Hospitality Management*, Vol. 95, p. 102930.

Liu, C. and Hung, K. (2022), "Improved or decreased? Customer experience with self-service technology versus human service in hotels in China", *Journal of Hospitality Marketing and Management*, Vol. 31 No. 2, pp. 176-204.

Lu, W. and Stepchenkova, S. (2015), "User-generated content as a research mode in tourism and hospitality applications: topics, methods, and software", *Journal of Hospitality Marketing & Management*, Vol. 24 No. 2, pp. 119-154.

Luo, H. and Yang, Y. (2016), "Intra-metropolitan location choice of star-rated and non-rated budget hotels: the role of agglomeration economies", *International Journal of Hospitality Management*, Vol. 59, pp. 72-83.

McGrath, M., Sherry, J. and Diamond, N. (2013), "Discordant retail brand ideology in the house of barbie", *Qualitative Market Research: An International Journal*, Vol. 16 No. 1, pp. 12-37, doi: 10.1108/13522751311289749.

Mcintosh, A.J. and Siggs, A. (2005), "An exploration of the experiential nature of boutique accommodation", *Journal of Travel Research*, Vol. 44 No. 1, pp. 74-81.

Massimo, D. and Ricci, F. (2019), "Clustering users' pois visit trajectories for next-poi recommendation", *Information and Communication Technologies in Tourism 2019*, Springer, Cham, pp. 3-14.

Meuter, M.L., Ostrom, A.L., Roundtree, R.I. and Bitner, M.J. (2000), "Self-service technologies: understanding customer satisfaction with technology-based service encounters", *Journal of Marketing*, Vol. 64 No. 3, pp. 50-64.

Mohd Dali, N. R. S., Abdul Hamid, H., Wan Nawang, W.R. and Wan Mohamed Nazarie, W.N.F. (2020), "Post pandemic consumer behavior: conceptual framework", *The Journal of Muamalat and Islamic Finance Research.*

Mowen, J.C., Wien, J.L. and Joag, S. (1987), "An information integration analysis of how trust and expertise combine to influence source credibility and persuasion" *ACR North American Advances*, Vol. 14 No. 564.

Nash, R., Thyne, M. and Davies, S. (2006), "An investigation into customer satisfaction levels in the budget accommodation sector in Scotland: a case study of backpacker tourists and the Scottish youth hostels association", *Tourism Management*, Vol. 27 No. 3, pp. 525-532.

Nguyen, N. and Leblanc, G. (2002), "Contact personnel, physical environment and the perceived corporate image of intangible services by new clients", *International Journal of Service Industry Management*, Vol. 13 No. 3.

Nguyen, T.H.D. and Vu, D.C. (2020), "Impacts of the COVID-19 pandemic upon mental health: perspectives from Vietnam", *Psychological Trauma: Theory, Research, Practice, and Policy*, Vol. 12 No. 5, pp. 480-481, doi: 10.1037/tra0000694.

Nowell, L.S., Norris, J.M., White, D.E. and Moules, N.J. (2017), "Thematic analysis: striving to meet the trustworthiness criteria", *International Journal of Qualitative Methods*, Vol. 16 No. 1, doi: 10.1177/1609406917733847.

Nozar, R. (2001), "Location is king, but perceived value has impact", *Hotel and Motel Management*, Vol. 216 No. 17, p. 8.

Pan, C.M. (2005), "Market structure and profitability in the international tourist hotel industry", *Tourism Management*, Vol. 26 No. 6, pp. 845-850.

Parasuraman, A., Zeithaml, V.A. and Berry, L. (1988), "SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.

Peña Ruiz, A. (2020), *An Approach to Designing an Online Teaching Proposal to Improve New Vocabulary Assimilation and Integration in English During the Covid-19 Pandemic*, available at: http://hdl.handle.net/10234/190182

Peng, J., Zhao, X. and Mattila, A.S. (2015), "Improving service management in budget hotels", *International Journal of Hospitality Management*, Vol. 49, pp. 139-148.

Philander, K. and Zhong, Y. (2016), "Twitter sentiment analysis: capturing sentiment from integrated resort tweets", *International Journal of Hospitality Management*, Vol. 55, pp. 16-24.

Pillai, S.G., Haldorai, K., Seo, W.S. and Kim, W.G. (2021), "COVID-19 and hospitality 5.0: redefining hospitality operations", *International Journal of Hospitality Management*, Vol. 94, p. 102869.

Rahimi, R. and Kozak, M. (2017), "Impact of customer relationship management on customer satisfaction: the case of a budget hotel chain", *Journal of Travel and Tourism Marketing*, Vol. 34 No. 1, pp. 40-51.

Ren, L., Qiu, H., Wang, P. and Lin, P.M. (2016), "Exploring customer experience with budget hotels: dimensionality and satisfaction", *International Journal of Hospitality Management*, Vol. 52, pp. 13-23.

Ruetz, D. and Marvel, M. (2011), "Budget hotels: low cost concepts in the US", *Europe and Asia. Trends and Issues in Global Tourism*, Vol. 2011, pp. 99-124.

Ruiz-Molina, M.E., Servera-Francés, D., Arteaga-Moreno, F. and Gil-Saura, I. (2018), "Development and validation of a formative scale of technological advancement in hotels from the guest perspective", *Journal of Hospitality and Tourism Technology*, Vol. 9 No. 3,

Rust, R.T. and Espinoza, F. (2006), "How technology advances influence business research and marketing strategy", *Journal of Business Research*, Vol. 59 Nos 10/11, pp. 1072-1078.

Ryu, K., Lee, H. R. and Kim, W. G. (2012), "The influence of the quality of the physical environment, food, and service on restaurant image, customer perceived value, customer satisfaction, and behavioral intentions", *International Journal of Contemporary Hospitality Management.*

Sabiote-Ortiz, C.M., Frías-Jamilena, D.M. and Castañeda-García, J.A. (2016), "Overall perceived value of a tourism service delivered via different media: a cross-cultural perspective", *Journal of Travel Research*, Vol. 55 No. 1, pp. 34-51.

Schmitt, B. (1999), "Experiential marketing", *Journal of Marketing Management*, Vol. 15 Nos 1/3, pp. 53-67.

Schmitt, B., Brakus, J.J. and Zarantonello, L. (2015), "From experiential psychology to consumer experience", *Journal of Consumer Psychology*, Vol. 25 No. 1, pp. 166-171.

Seyitoğlu, F. and Ivanov, S. (2020), "A conceptual framework of the service delivery system design for hospitality firms in the (post-) viral world: the role of service robots", *International Journal of Hospitality Management*, Vol. 91, p. 102661.

Shen, S., Xu, K., Sotiriadis, M. and Wang, Y. (2022), "Exploring the factors influencing the adoption and usage of augmented reality and virtual reality applications in tourism education within the context of COVID-19 pandemic", *Journal of Hospitality, Leisure, Sport and Tourism Education*, Vol. 30, p. 100373.

Shin, H. and Kang, J. (2020), "Reducing perceived health risk to attract hotel customers in the COVID-19 pandemic era: Focused on technology innovation for social distancing and cleanliness", *International Journal of Hospitality Management*, Vol. 91, p. 102664.

Sigala, M. (2020), "Tourism and COVID-19: impacts and implications for advancing and resetting industry and research", *Journal of Business Research*, Vol. 117, pp. 312-321.

Sigala, M. (2015), "Gamification for crowdsourcing marketing practices: applications and benefits in tourism", *Advances in Crowdsourcing*, pp. 129-145.

So, K.K.F. and King, C. (2010), "When experience matters": building and measuring hotel brand equity: the customers' perspective", *International Journal of Contemporary Hospitality Management*, Vol. 22 No. 5.

Tanford, S., Raab, C. and Kim, Y.S. (2012), "Determinants of customer loyalty and purchasing behavior for full-service and limited-service hotels", *International Journal of Hospitality Management*, Vol. 31 No. 2, pp. 319-328.

Thomas, M., Chun, H., Gupta, S., Liaukonyte, J., Shu, S. and Woolley, K. (2022), "Authentically cornell: the evolution of marketing at cornell", *Customer Needs and Solutions*, pp. 1-6.

Titz, K. (2009), "Experiential consumption: affect–emotions–hedonism", *Handbook of Hospitality Marketing Management*, Routledge, pp. 346-374.

Turpo Gebera, O. W. (2008), "La netnografía: un método de investigación en Internet", Educar.

Walker, J.R. (2008), *Exploring the Hospitality Industry*, 2/e. Pearson Education India.

Walls, A., Okumus, F., Wang, Y. and Kwun, D.J.W. (2011), "Understanding the consumer experience: an exploratory study of luxury hotels", *Journal of Hospitality Marketing and Management*, Vol. 20 No. 2, pp. 166-197.

Walther, C.S. and Sandlin, J.A. (2011), "Green Capital and social reproduction within families practicing voluntary simplicity in the US", *International Journal of Consumer Studies*, Vol. 37 No. 1, pp. 36-45.

Wilkins, H., Merrilees, B. and Herington, C. (2007), "Towards an understanding of total service quality in hotels", *International Journal of Hospitality Management*, Vol. 26 No. 4, pp. 840-853.

Wong, I.A. (2013), "Exploring customer equity and the role of service experience in the casino service encounter", *International Journal of Hospitality Management*, Vol. 32, pp. 91-101.

Wu, C.H.J. (2007), "The impact of customer-to-customer interaction and customer homogeneity on customer satisfaction in tourism service – the service encounter prospective", *Tourism Management*, Vol. 28 No. 6, pp. 1518-1528.

Wu, S.H. and Gao, Y. (2019), "Understanding emotional customer experience and co-creation behaviours in luxury hotels", *International Journal of Contemporary Hospitality Management*, Vol. 31 No. 11.

Xiang, Z. and Krawczyk, M. (2016), "What does hotel location mean for the online consumer? Text analytics using online reviews", *Information and Communication Technologies in Tourism 2016*, Springer, Cham, pp. 383-395.

Yavas, U. and Babakus, E. (2005), "Dimensions of hotel choice criteria: congruence between business and leisure travelers", *International Journal of Hospitality Management*, Vol. 24 No. 3, pp. 359-367.

Yuan, Y.H.E. and Wu, C.K. (2008), "Relationships among experiential marketing, experiential value, and customer satisfaction", *Journal of Hospitality and Tourism Research*, Vol. 32 No. 3, pp. 387-410.

Zeithaml, V.A. (1988), "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence", *Journal of Marketing*, Vol. 52 No. 3, pp. 2-22.

Zeithaml, V.A., Parasuraman, A. and Berry, L.L. (1985), "Problems and strategies in services marketing", *Journal of Marketing*, Vol. 49 No. 2, pp. 33-46.

Zeleti, F.A., Ojo, A. and Curry, E. (2016), "Exploring the economic value of open government data", *Government Information Quarterly*, Vol. 33 No. 3, pp. 535-551.

Zhang, Y. and Hitchcock, M.J. (2017), "The Chinese female tourist gaze: a netnography of young women's blogs on Macao", *Current Issues in Tourism*, Vol. 20 No. 3, pp. 315-330.

Zhang, Y., Xiong, Y. and Lee, T.J. (2020), "A culture-oriented model of consumers' hedonic experiences in luxury hotels", *Journal of Hospitality and Tourism Management*, Vol. 45, pp. 399-409.

Zhang, Y., Yuan, Y., Su, J. and Xiao, Y. (2021), "The effect of employees' politeness strategy and customer membership on customers' perception of co-recovery and online post-recovery satisfaction", *Journal of Retailing and Consumer Services*, Vol. 63, p. 102740.

Web reference

Available at: https://unctad.org/system/files/official-document/ditcinf2021d3_en_0.pdf (accessed 12 January 2022).

Corresponding author
Devika Rani Sharma can be contacted at: devikayekkelli@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835831

Feature Ranking using Statistical Techniques for Computer Networks Intrusion Detection

Yash Sharma
Department of Applied Mathematics
Delhi Technological University
Delhi-110042, India.
yash9013562595@gmail.com

Somya Sharma
Department of Applied Mathematics
Delhi Technological University
Delhi-110042, India.
contactsomya26@gmail.com

Anshul Arora
Department of Applied Mathematics
Delhi Technological University
Delhi-110042, India.
anshul15arora@gmail.com

***Abstract*—Nowadays with the enlargement of the utility of computer networks, the danger of getting our personal information being misused has also grown rapidly. According to the [1] 2022 Cyber Threat Report released by SonicWall, the health care industry faced a 755% increase in the attacks worldwide and Governments witnessed a 1,885% hike in attacks in 2021. Moreover, the threats due to intrusions are becoming more and more complex and difficult to detect. This research study intends to observe and inspect the network traffic of normal as well as intrusions and to perceive the network traffic features that can distinguish normal traffic from intrusions traffic. In order to extract distinguishing features, the proposed research study utilizes the statistical techniques of ANOVA and the Chi-Square test to order and line up the network traffic features. The experimental results demonstrate that the feature Bytes Sent is ranked at the top when the ANOVA test is applied over the data. In the cases where Chi-Square test is applied over normal and malware data, the features bytes sent and bytes received are ranked at the top respectively. Such ranking can help build an effective intrusion detection system.**

***Index Terms*—Feature Ranking, Intrusion Detection, Network Traffic, Traffic Analysis.**

I. Introduction

Cyber-attacks have been on an exponential rise, not only in terms of numbers but also in terms of diversity, in the past few years with the constant growth in the usage of the Internet. Malicious activities, also sometimes called malware / intrusions attacks, have escalated to a great extent in the past few years like never before. These attacks, including the zero-day exploits have gained media coverage as well due to the threats posed by them to the user community. In the modern era of threats, computer networks have gained a very important security focus over the past few decades. According to the [1] 2022 Cyber Threat Report released by SonicWall, the health care industry witnessed a 755% hike in the attacks worldwide and Governments witnessed a 1,885% hike in attacks in 2021. Cyberspace security can be defined as a collection of procedures that are planned to guard or protect systems, networks and personal files or folders from cyber threat, unlawful approach, modification or damage whereas intrusions can be referred as endeavours to compromise the secrecy availability of a structure.

An Intrusion Detection System (IDS) is one of the most crucial measure meant to provide the security to the system against any malicious activity. It provides the required security by continuously monitoring the system attributes such as Internet traffic, also called as network traffic. Hence, the research community defines "Intrusion Detection" as the process of continuously monitoring the network traffic or system calls or other similar events to detect any malicious or unauthorized activity on the system. Any software, device or application whose sole purpose is to monitor cyber threats can be considered as an IDS. Well-constructed IDS can monitor network traffic and other activities within the entire network and not just at its perimeter. But in contrast to firewalls, IDS cannot take any kind of action themselves and can only play a monitoring role. Therefore, they need an administrator to process their alerts.

To identify intrusions effectively and efficiently in a constantly evolving environment is not an easy task. People usually are unable to produce adequate design decisions in such adverse conditions and moreover the profile of the intruders also keeps on changing with time. Risks have been multiplying day by day and they come in a variety of patterns and ways which makes it pretty hard for security groups to do their job with the constant changing faces of more threats, attacks and responses using right countermeasures. The tools that earlier proved to be effective such as firewalls, different policies to secure access control and encryption have proven to be inefficient in protecting networks and systems in the right manner in this age. This kind of situation requires far better and more intelligent countermeasures to keep intact the security of networks and important systems, and that can be provided by IDS. IDS can be grouped or we can say placed into three categories according to the detection method they use namely signature-based detection, anomaly-based detection, and hybrid detection.

- Signature based detection approach, also known by the name of "misuse detection" maintains a collection or database of signatures of known attacks. While monitoring the data to detect any malicious activity, the model compares the data with the signatures present in the database. This kind of a system checks the input system for the presence of an attack pattern. Such signature-based database must be updated at regular intervals to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:39 UTC from IEEE Xplore. Restrictions apply.

detect newer/zero-day attacks. However, even on adding the newer signatures to the database regularly, it can't be said that this system is supremely efficient.

- Anomaly detection tries to study the "normal" behaviour of the system. Any deviation from this expected behaviour will be considered as a threat. This method is somewhat better as it does not require updates or even the presence of a database.
- Hybrid detection combines the two principles of the above methods to mitigate weaknesses of each of them. Such type of detection either first employs anomaly detection technique followed by misuse detection approach or first misuse detection technique followed by anomaly detection approach. The objective is to detect known attacks/intrusions and unknown attacks/intrusions using the signature based and anomaly based detection techniques respectively.

What makes an effective security system is a robust Intrusion Detection system (IDS) that can deal properly with all sort of attacks and threats before any damage. The main features that an IDS must have are effectiveness, adaptability and extensibility. Effectiveness can be defined as the potential to detect large as well as comparatively small attacks whereas extensibility refers to as the ability of an IDS to tailor itself to various environments adapt to them easily. If we take a closer look at the previous works in this field, one might notice that most of the work carried out is related to signature detection, but despite those works being effective, they raise two important issues that must be further looked at. A human is responsible for the making and maintenance of signatures which in turn raises the issue of cost and affording labour. Other issues include the time lag between updating the signatures the hard way and dealing with the new attacks and moreover they don't work with attacks that haven't been updated yet.

Contributions : The main contributions of the paper are summarized below:

1) In this paper, the network traffic features are extracted from the traffic files of normal and intrusion traffic.
2) We apply the statistical techniques of the ANOVA and Chi-Square Test on the set of traffic features intending to rank the traffic features. We obtain the ranking of features with these statistical tests. Such feature ranking will help in building effective intrusion detection systems based upon top-ranked features only.

The remainder of the paper is designed as follows. We review the related work in the field of intrusion detection in Section II. We discuss the techniques for feature ranking in Section III, and report the results obtained from proposed techniques in Section IV. And finally, we conclude with future work directions in Section V.

II. RELATED WORK

This segment comprises of the related work in our associated filed. Some of the works such as [20] - [27] have been proposed in the literature to detect intrusions in mobile environment. However, we do not aim to build IDS for mobile networks. Therefore, we discuss the related work in the field of desktop based intrusion detection. Authors in [2] talked about network malware detection being of utmost use as it is capable of keeping a track of the stream of unknown data coming in the network as well as creating a warning-based system as a precautionary measure. This system which puts a hault over trespassing was based on FCMA and was completely capable of modifying the detection system and sorting the traffic flow. Further, they build a model by applying the Machine Learning algorithms. This paper gave us an insight about the methods of identification and selection of features in the field.

The authors in [3] proposed an IP flow-based (IDS) system to monitor and defend IOT from all kinds of dangers. The idea discussed in this paper is to collect the IP flows, where IP stands for "Internet Protocol", from an IoT network to monitor them in order to keep a check over all kinds of malicious activities in the given related field. This approach was meant to work for both IOT as well as its properties and included both centralized examining as well as distributed content collecting components. The authors concluded the paper by stating that their proposed idea is capable of achieving complete success with no wrong detection. Dasari et al. [4] used meta-heuristic association to bring out the least door value for the affair, followed by making use of a classifier to interpret the affair is benign/threat. The corresponding property used for the paper is drift detection. Binbusayyis et al. [5] talked about the growing need of an intrusion detection system and cybersecurity as a primary defense mechanism that can learn as well as protect the computing things of value from the never ending threats. The idea behind this paper is somehow related to deep learning that too unsupervised yet unlike other methods present in the market.

Dilemma of keeping an eye on malware in MANET's was taken up in [6] followed by a never been used. The system proposed used time variant snapshots to dig out routing attacks. In [7], novel intrusion detection technique is proposed with a fully automatic attack signature generation capability. The technique depicted in this aper can be utilised to detect possible incoming of malwares. authors in [8] introduced an easy to use with very little load detection method with great efficiency for IOT networks.

Multi architectural modular deep neural network system has been proposed to bring down the wrong detection of malware and threats by Atefni et al. [9]. The experimental results showed a crazy response showing success as high as possible when kept in balanced with the ongoing works . Gupta et al. [10] talked about a multi behaviour fusion-based IDS where stream data mining is done by using ST-SR (stochastic relaxation). The major contribution of this paper is that it can easily predict the quality or the standard of utility of each network amidst the flow and also allowed the enjoyer to keep in touch with a conserved and reliable web system.

A deep multi-scale convolution neural network is being proposed by Wang et al. [11] for network intrusion detection. The author used two different testing sets of which the task to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:39 UTC from IEEE Xplore. Restrictions apply.

classify the second one was a bit hader. Out-turn have shown the accuracy rate to reach up to 98%. The major help done by this paper was that it can overcome the limitation of using traditional detection methods. Hsu et al. [12] used a deep learning scheme to overcome difficulties faced by applying machine learning to IDS. The work used in this paper is related to neural network layers CNN which can be used to sort every binary track data. This data set included two testing sets. The work proposed in [13] is based on univariant ensemble feature reduction and choosing techniques that are very well capable of reducing the dataset for better working and efficiency. The authors in [14] built a classified system which works upon tree-based codes. The experimental values showed detection accuracy more than 80% which was indicated by making use of a couple of ML algorithms related to the sum rule. It proved to be better than individual random tree algorithm.

Caligari et al. [15] used an approach using different sorts of Markovian structures which can be characterized as a statistical approach that was fully capable of providing an edge over other statistical technique used before. Y.Zhu et al. [16] introduced a model related to unsupervised clustering, that worked as a combination of Deep Auto-Encoder(DAE) and Gaussian Mixture Model(GMM) which adopted a two-layer hierarchial form.

M. Dhingra et al. [17] worked with the motive of applying numerous machine learning techniques on a particular dataset and to note down the detection results. A well known pre-processed dataset has been used and in the end the relevant 12 features are put under machine learning algorithms to produce significant detection accuracy. J.Zhou et al. [18] utilised the C4.5 decision tree algorithm to take out a variety of network attacks from a preprocessed data. According to the results the model is able to produce good accuracy rate as well as lower false positive rate. S. Ding et al. [19] combined CNN(Convolutional Neural Networks) and BiSRU (Bi-directional Simple Recurrent Unit) and in the end they output the categorization outcomes by softmax.

However, in this paper, different from other works, we focused on ranking the traffic features which would enable us to find out the best distinguishing features that can better classify normal and malicious /intrusion traffic. We discuss the ranking techniques in the detailed manner the next section.

III. FEATURE RANKING

In this section, we explain our proposed methodology in detail. Fig. 1. gives a brief yet complete idea of our approach. The following subsections discuss different phases of the proposed work.

A. Dataset Collection

Firstly, we intend to gather the normal and intrusion network traffic for further use in our experiments. We accumulate the normal traffic on our systems connected to the Internet using Wireshark. Wireshark captures the traffic in the form of pcap files. Further, we got the intrusion traffic from Canadian Institute for Cybersecurity. Similar to benign traffic, the intrusion traffic was also in the form of pcap files that could easily be read by Wireshark. To extract the network traffic features , we utilised TCP flows.

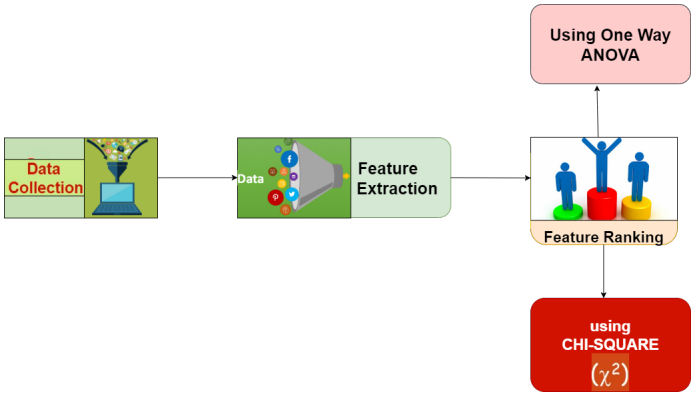


Fig. 1. Proposed Methodology

B. Network Traffic Features

After collecting data, we extracted some network traffic features from the normal and intrusion traffic. Table I outline the traffic features that we have used in our experiments. All such features are flow-based and can easily be extracted from the pcap flows of normal and intrusion traffic. For each feature, we have also written its short notation, for example F2 for Time interval between packets sent, F4 for flow duration, etc. Later in the results section, we represent the features by their notations for easy understanding and interpretation of the results.

TABLE I
LIST OF NETWORK TRAFFIC FEATURES

Average Packet Size **(F1)**	Time interval between packets sent **(F2)**	Time interval between packets received **(F3)**
Flow Duration **(F4)**	Ratio between Incoming to outgoing packets **(F5)**	Ratio of Incoming to outgoing bytes **(F6)**
Packet size sent **(F7)**	Packet size received **(F8)**	Bytes sent **(F9)**
Bytes received **(F10)**	Number of packets sent **(F11)**	Number of packets received **(F12)**

C. Features Ranking

Statistical tests were used to rank the features we extracted. We aim to rank the features since it helps in eliminating the irrelevant features and their removal will help in improving the detection accuracy. Hence, with this goal, we rank the features using two statistical tests, namely ANOVA and Chi-Square Test. We discuss each of them following:

1) ANOVA Test:
ANOVA stands for Analysis of Variances between the two populations/data we have. When the comparison of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:39 UTC from IEEE Xplore. Restrictions apply.

two or more than two groups/data is required, oneway analysis of variance, i.e., One-way ANOVA is a suitable method. We can say that this method classify the relative size of variants between groups called group variance compared to the average variance within the group called within-group variance.As the name suggests, One Way(one factor), it is a clear indication that there exists only single variable("treatment") but with multiple levels and only a single level of treatment works at any given time for a given subject. It is a statistical method that verifies the impact of several factors by comparing the average or means of various sample data. ANOVA figures out if the groups or classes formed by the level of an independent variable are statistically similar or not. The ANOVA method determines this by evaluating the means of the treatment levels, and it further labels whether the means are distinct from the overall mean of the dependent variable or not. It is a statistical method and hence like all of them it also works on the principle of hypothesis which is a null and alternate hypothesis. The null hypothesis in ANOVA is valid when the averages of all sample data are equal or we can say that they don't have any remarkable difference between them. The alternate hypothesis in ANOVA is valid when at least one of the sample data means varies from the other values in the sample data.

2) Chi-Square Test:
The major application of this test is to determine if the variables of different categories defined are independent of each other or not. It can also be defined as if there is any significant difference between variables and their expected values. The Chi-square formula is defined in the below equation.

$$\chi_c{}^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where:
c=Degrees of freedom
O=Observed value(s)
E=Expected value(s)

According to the null hypothesis of this test there is no link between the original and expected data. The alternate hypothesis says that the original and expected data are dependent on each other.
Using the above two techniques, we rank the features. Chisquare test is applied separately on normal and intrusion files, whereas, ANOVA test is applied by merging the normal and intrusion traffic. Hence, we get one ranking from the ANOVA test and two rankings from the Chi-Square test.

IV. RESULTS

In this section, we report the results obtained, i.e., we study the ranking of features obtained with ANOVA and Chi-Square test.

A. Features Ranking

When applied ANOVA tests, we get the the p-value as well for each feature (Table II). The P-value obtained is directly proportional to rank of feature. Therefore, the features are placed in a descending order from top to bottom, where top most feature is the best ranked according to the ANOVA test, i.e., F9 and the last feature is the least in ranking which is F6.

TABLE II
RANKING USING ANOVA TEST

S. No.	ANOVA RANKING	ANOVA TEST VALUE : P VALUE
1	F9	4.30E -72
2	F2	0
3	F3	0
4	F5	3E-128
5	F7	0
6	F11	8E-147
7	F10	0
8	F8	1.5E-18
9	F12	0.12025
10	F1	7.3E-12
11	F4	1E-11
12	F6	5.7E-21

On applying chi-square test on normal data and the malware data separately we get their respective chi square value of each feature as follows (Table III): We found that higher the chisquare value for the feature higher it is ranked. Therefore, F9 is the highest ranked feature and F5 is the least ranked feature in normal data. And F10 is the highest ranked feature and F5 is the least ranked feature in malware data.

TABLE III
RANKING USING CHI SQUARE TEST(NORMAL)

S. No.	CHI-SQUARE NORMAL RANKING	CHI-SQUARE NORMAL TEST STATIC VALUE
1	F9	2122715.611
2	F10	37495.33308
3	F12	40352.17939
4	F8	19016927.82
5	F11	4445.849973
6	F1	339624.5391
7	F4	541806.182
8	F7	3249576.305
9	F6	5.12082E+11
10	F3	1.94795E+11
11	F2	3081397.304
12	F5	111805954.4

Table II and Table III highlights the ranking of features obtained from ANOVA test and Chi square test on normal data respectively. Here we observe that F9 feature, i.e., Bytes Sent is the top-ranked feature in ANOVA as well as Chi-Square ranking (on normal).
Whereas F10 feature, i.e., Bytes Received, is the top-ranked feature in Chi-Square ranking on malware in Table IV . F5 is the least ranked feature in the Chi-Square test (on both malware and normal), and F6 is the least ranked one in the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:39 UTC from IEEE Xplore. Restrictions apply.

ANOVA ranking. We witness that we get the different ranking of features across ANOVA and Chi-Square tests.

TABLE IV
RANKING USING CHI SQUARE TEST(MALWARE)

S. No.	CHI-SQUARE MALWARE RANKING	CHI-SQUARE MALWARE TEST STATIC VALUE
1	F10	5926304.712
2	F9	81805.3058
3	F12	70452.7675
4	F11	581449.6579
5	F8	24584.32825
6	F1	209980.1579
7	F7	1052382.224
8	F4	18539138.24
9	F6	3401902043
10	F2	56885444238
11	F3	23331382.89
12	F5	31027067.28

In the future, we can use such feature ranking for effective intrusion detection.

V. CONCLUSION AND FUTURE WORK

We have proposed methods of ranking features using statistical tests. We used ANOVA and Chi Square test to rank different features extracted from our network traffic of both normal and intrusion type. Feature F9 and F10 which are Bytes sent and Bytes received happens to be best amongst all. This method of ranking can be further used for deriving best set of features that can detect intrusion more accurately. For our possible future work, we aim to use this ranked feature list to create an efficient Intrusion Detection system.

REFERENCES

[1] https://fortune.com/2022/02/17/ransomware-attacks-surge-2021-report/: :text=Governments%20worldwide%20saw%20a%201%2C885, SonicWall%2C%20an%20internet%20cybersecurity%20company.

[2] H.Li , "Research on network intrusion detection technology based on improved FCMA algorithm," Journal of Ambient Intelligence and Humanized Computing, 2021.

[3] L. Santos, R. Gonc¸alves C.Rabadao, and J.Martins, " A flow-based intrusion detection framework for internet of things networks," in SPRINGER Cluster Computing, 2019.

[4] D. B.Dasari, G. Edamadaka, Ch. S. Chowdary, and M. Sobhana, "Anomaly based network intrusion detection with ensemble classifiers and meta heuristic scale (ECMHS) in traffic flow streams," Journal of Ambient Intelligence and Humanized Computing , vol. 12, pp. 9241-9268, 2021.

[5] A. Binbusayyis, and T. Vaiyapuri, "Unsupervised deep learning approach for network intrusion detection combining convolutional autoencoder and one-class SVM," in Springer Applied Intelligence, vol. 51, pp. 7094-7108,2021.

[6] K. Bala, S. Jothi, A. Chandraseka, "An enhanced intrusion detection system for mobile ad-hoc network based on traffic analysis," 123 Cluster Computing (2019) vol. 22, pp. S15205–S15212, 2019.

[7] A Boulaiche,and K Adi, " An auto-learning approach for network intrusion detection," Telecommunication System, vol. 68, pp. 277–294, 2018.

[8] A. Davahli, M. Shamsi, and G. Abaei, " Hybridizing genetic algorithm and grey wolf optimizer to advance an intelligent and lightweight intrusion detection system for IoT wireless networks," journal of Ambient Intelligence and Humanized Computing vol. 11,pp. 5581–5609,2020.

[9] R. Atefni ,and M. Ahmadi, "Network intrusion detection using multi architectural modular deep neural network," The Journal of Supercomputing vol. 77,pp. 3571–3593,2021.

[10] A. R.Gupta ,J. Agrawal, "The multi demeanor fusion based robust intrusion detection system for anomaly and misuse detection in computer networks," journal of Ambient Intelligence and Humanized Computing vol. 12,pp. 303–319 , 2021.

[11] X. Wang, S. Yin1, H. Li, J. Wang,and L. Ten, "A Network Intrusion Detection Method Based on Deep Multi scale Convolutional Neural Network," International Journal of Wireless Information Networks vol. 27,pp. 503–517, 2020.

[12] C.M. Hsu, M. Z Azhari, H.Y.Hsieh, S. W.Prakosa,and J.S Leu, "Robust Network Intrusion Detection Scheme Using Long-Short Term Memory Based Convolutional Neural Networks", Mobile Networks and Applications, 2020.

[13] S. Krishnaveni,S. Sivamohan, S. S. Sridhar, and S. Prabakaran, "Efficient feature selection and classification through ensemble method for network intrusion detection on cloud computing," 123 Cluster Computing, vol. 24, pp. 1-19, 2021.

[14] Kevric, S. Jukic and A. Subas, "An effective combining classifier approach using tree algorithms or network intrusion detection" , Neural Computer and Application, vol. 28 (Suppl 1),pp. S1051–S1058, 2017.

[15] C. Caligari, S. Vaton, and M. Pagano, "Communication Networks A new statistical method for detecting network anomalies in TCP traffic", EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS Eur. Trans. Telecoms, vol. 21, pp. 575–588 ,15 June 2010.

[16] Y. Zhu, D. Han, and X. Yin, "A hierarchical network intrusion detection model based on unsupervised clustering," Proceedings of the 13th International Conference on Management of Digital EcoSystemsNovember, 2021,pp. 22–29.

[17] M.Dhingra, S.C.Jain, R.S. Jadon, "Performance of Machine Learning Approaches for Malicious Traffic Intrusion Detection in Network," Advances in Clean Energy Technologies . Springer Proceedings in Energy. Springer, Singapore(2021).

[18] J. Zhou, X. Jiang, C. Liu, J. Zhang, L. Liao and J. Lu, "Multi-Traffic Features Network Intrusion Detection Algorithm Based on C4.5," 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), 2021, pp. 548-552,2021.

[19] S. Ding, Y. Wang and L. Kou, "Network intrusion detection based on BiSRU and CNN," 2021 IEEE 18th International Conference on Mobile Ad Hoc and Smart Systems (MASS), pp. 145-147,2021.

[20] A. Arora, S. Garg and S. K. Peddoju, "Malware Detection Using Network Traffic Analysis in Android Based Mobile Devices," 8th International Conference on Next Generation Mobile Apps, Services and Technologies, pp. 66-71, 2014.

[21] A. Arora, S. K. Peddoju and M. Conti,"PermPair: Android Malware Detection Using Permission Pairs," in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 1968-1982, 2020.

[22] A. Arora and S. K. Peddoju, "Minimizing Network Traffic Features for Android Mobile Malware Detection', Proceedings of the 18th International Conference on Distributed Computing and Networking, Article 32, pp. 1–10, 2017.

[23] A. Arora and S. K. Peddoju, "NTPDroid: A Hybrid Android Malware Detector Using Network Traffic and System Permissions," 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), pp. 808-813, 2018.

[24] K. Khariwal, J. Singh and A. Arora, "IPDroid: Android Malware Detection using Intents and Permissions," 4th World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), pp. 197-202, 2020.

[25] A. Arora, S. K. Peddoju, V. Chouhan, and A. Chaudhary, "Hybrid Android Malware Detection by Combining Supervised and Unsupervised Learning," In Proceedings of the 24th Annual International Conference on Mobile Computing and Networking, pp. 798–800, 2018.

[26] M. Upadhayay, A. Sharma, G. Garg and A. Arora, "RPNDroid: Android Malware Detection using Ranked Permissions and Network Traffic," 5th World Conference on Smart Trends in Systems Security and Sustainability (WorldS4), pp. 19-24, 2021.

[27] G. Garg, A. Sharma, and A. Arora, "SFDroid: Android Malware Detection Using Ranked Static Features", in International Journal of Recent Technology and Engineering, vol. 10, no.1, pp. 142-152, 2021

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:39 UTC from IEEE Xplore. Restrictions apply.

Biomedical Microdevices (2022) 24:29
https://doi.org/10.1007/s10544-022-00630-2

Graphene and graphene nanohybrid composites-based electrodes for physiological sensing applications

Bani Gandhi[1] · **Nallanthighal Srinivasa Raghava**[1]

Accepted: 31 July 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
In this paper, three categories of ECG electrodes were fabricated. Graphene/PDMS(Polydimethylsiloxane)(G-I), Graphene/MWCNT-COOH(Carboxylic-acid functionalized Multi-walled Carbon Nanotubes)/PDMS(G-II),and Graphene/SWCNT-COOH(Carboxylic-acid functionalized Single-walled Carbon Nanotubes)/PDMS(G-III). Each group had thirteen electrodes with varying concentrations ranging from 0.1-5wt%. Since CNTs get tangled easily, it becomes necessary to disperse them properly. To achieve optimal dispersion, CNTs were first sonicated with Isopropyl Alcohol (IPA), and then with PDMS. Mold casting was the technique used for fabricating the electrodes. The results were compared with the conventional ECG electrodes. Best results were achieved from G-III at 3wt% as the value of capacitance is high (0.172nF) as compared to G-I and G-III values at 3wt% which are 0.036nF (0.036nF) and 0.015nF respectively. As capacitance has an inverse relationship with the resistance and impedance, thus at 3wt% the resistance (0.361MΩ) and impedance (0.36MΩ) values are low, which satisfies the relationship. The values of resistance and impedance of G-II are low when compared with the values of G-I and G-II. Great results and ECG waveform are achieved with 3wt% for G-II, which also uses less nanomaterials to produce such great ECG results. It was observed that even after using the electrodes for 5 days, the ECG signal did not degrade over time and no skin allergies were detected for any of the three groups. The ECG tracking system was developed on the concept of the Internet-of-Things (IoT) using various electronic hardware components and software solutions. The results from the fabricated electrodes were promising and were suitable for long-term, and continuous ECG monitoring.

Keywords Graphene · MWCNT-COOH · SWCNT-COOH · PDMS · Electrocardiogram (ECG) · Internet-of-Things (IoT)

1 Introduction

As the prevalence of heart diseases, hypertension, and other non-communicable diseases (NCD) in the world increases, there will be an excessive need and desire for personal and individual healthcare devices which can sense and perceive a person's ECG at any remote location and time. The evolution of self-health monitoring systems accelerates medical assistance not only for elderly and feeble people but also aids the concept of real-time monitoring. This averts grievous damage to patients going through various heart diseases. This can be made possible by using lightweight ECG electrodes fabricated using nanomaterials. Self-health monitoring can be achieved by enabling the person to wear nanomaterial-based ECG electrodes. It will help acquire the ECG signal, which can be sent to the medical expert with the help of various wireless communication protocols like Wi-Fi, 4G, etc. conversely, the expert can provide necessary preventive measures to the affected person. It is pictorially depicted in Fig. 1. The flow of the figure is such that the patient (Ruoff et al. 1993) is wearing Nanomaterial-based ECG electrodes/sensors which are lightweight, flexible, and portable (Lu 1997; Yu et al. 2000). The person can carry the sensors anywhere and everywhere. When the ECG signal is acquired by the sensor, it is sent to an aggregator which collects the data from the patient and links it to the server or the internet.

This can be done either by using IoT-based smart boards or wireless communication protocols. Subsequently, the data is saved on the server for analysis and assessment by the medical expert. Also, the data is transmitted to the doctor

✉ Nallanthighal Srinivasa Raghava
nsraghava@gmail.com

Bani Gandhi
bani.gandhi@gmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University (DTU), Delhi 110042, India

Published online: 23 August 2022

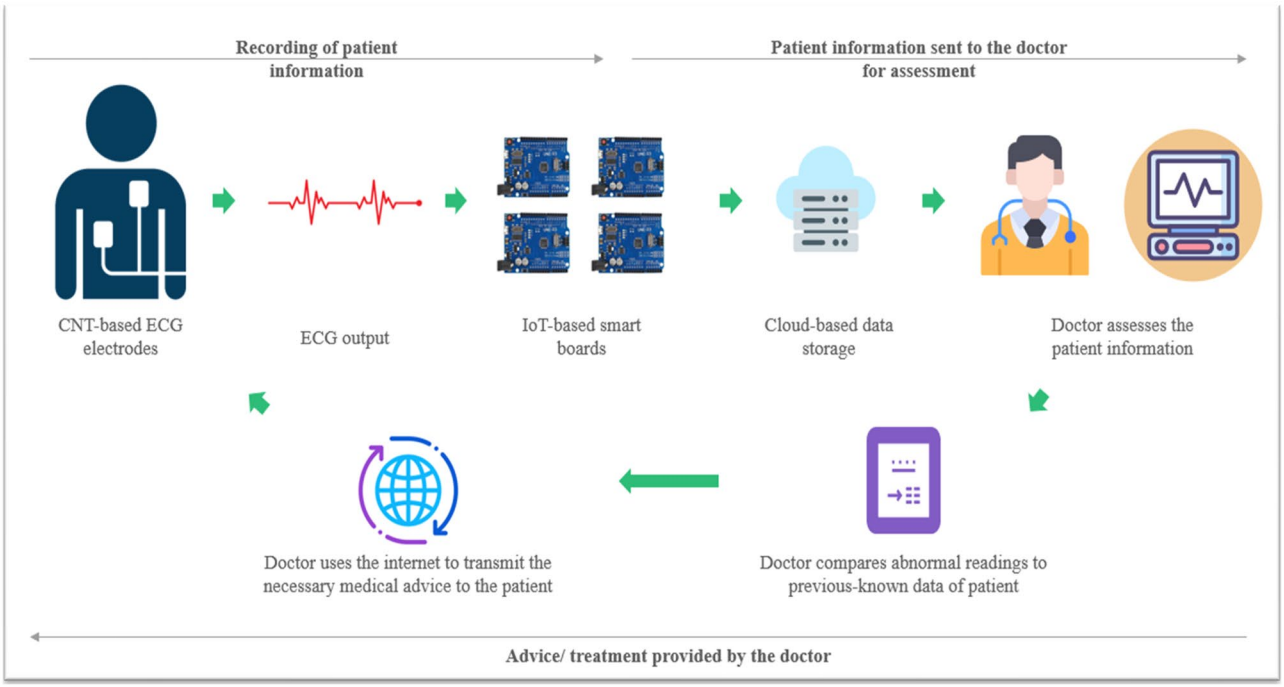


Fig. 1 SmarTele ECG monitoring system

so that immediate help can reach the patient suffering from heart disease. When the abnormal ECG reaches the doctor, he/she analyses the signal and detects what is wrong with it, and accordingly assists the patient with medical advice, treatment suggestions, or urgent care as needed.

In this paper, three categories of electrodes have been fabricated, Graphene/PDMS (G-I), Graphene/MWCNT-COOH/PDMS (G-II), and Graphene/SWCNT-COOH/PDMS (G-III). The ECG data from the person is acquired via AD8232. The collected data via Wi-Fi is transferred to the IoT cloud platform and saved to a database (MongoDB).

2 Experimental

2.1 Materials

1. Graphene, SWCNT-COOH, and MWCNT-COOH

   Empirical evidence suggests many ECG sensors have been fabricated, using nanomaterials (Graphene and CNTs). This is because a good orientation is observed while mixing nanomaterials (Graphene and CNTs) with polymers or placing them on the substrate. Further, adding to the advantages of nanomaterials they can be used for fabricating both functional and structural nanocomposites to be used for ECG monitoring (Ruoff et al. 1993; Yu et al. 2000). Also, the high surface-to-volume ratio of these materials permits high biomolecule loading per unit geometric area that leads to high signal amplification. Graphene (Thermo Fischer Scientific) was used as it can be manufactured in bulk at a low-cost, also it is easier to alter the characteristics of graphene, and possess an impressive aspect ratio (Zhu et al. 2017). Functionalized CNTs possess better-dispersing capabilities, which eventually help in attaining excellent interfacial bonding between the CNTs and the polymer. The bonding further increases the conductivity of the sensor (Nag et al. 2016). The functionalities were purchased from Nano Research Elements with specifications shown in Table 1.
2. Sylgard 184 Silicone Elastomer (PDMS)

   Polymer-based electrodes have fascinated growing attention for long-term medical tracking systems (Baek et al. 2008). These electrodes provide comfortable wearing and are a great substitute for the Ag/AgCl electrodes because of their disadvantages which lead to signal degradation over time and cause some allergic reactions (Pantelopoulos and Bourbakis 2010). Since CNTs are twisted and mustered randomly, which leads to better contact with the polymer composite. Thus, the fabricated electrode can be achieved with very low concentrations of CNTs (Lee et al. 2015).
3. Isopropyl Alcohol

   As one of the properties of CNTs is that they tend to form clusters which promote an increase in van der Waals force between the CNTs. This further disturbs

Table 1 Specifications of SWCNT-COOH and MWCNT-COOH

S.No	Parameter	Graphene	SWCNT-COOH	MWCNT-COOH
1.	Purity	99.9%	> 98wt%	> 98wt%
2.	Average diameter	1.5 µm	1–4 nm	10–30 nm
3.	Average Length	3 nm	0.8–1.6 nm	10–30 µm
4.	Special Surface Area	500 m^2 /g	5–30 m^2/g	180 m^2/g
5.	Electrical conductivity	0.3*10^6 S/m	> 100 S/cm	> 100 S/cm
6.	Tap Density	0.2–0.4 g/cm^3	690 m^2/g	0.27 g/cm^3

the solubility of CNTs in the solvent (Mallakpour and Soltanian 2016). Therefore, in this research, a two-step dispersion method is followed. CNTs are first dispersed in Isopropyl alcohol (IPA) and then later in PDMS, to avoid cluster formation. IPA was chosen as both CNTs and PDMS were soluble. As non-polar solvents like methanol swelled up the PDMS (Kim et al. 2018).

2.2 Preparation of the mold

The technique for casting electrodes was performed using molds. The process of making the molds is facile and unchallenging. Molds were formed using Sylgard 184 Silicone Elastomer (Dow Corning) base 184A and curing agent Sylgard 184B (The Dow Chemical Company 2017). A mix of 184A and 184B is prepared in the ratio of 10:1. This mixture is poured into the petri dish. The petri dish is prepared before pouring the PDMS mix. A double tape is used to stick the acrylic sheet (diameter:50 mm, thickness:15 mm). The acrylic sheet is used to form the cavity inside the mold for pouring the CNT/PDMS composite. The PDMS mix was then poured into the petri dish and placed in the oven for 30 min at 80 °C. The molds were separated from the petri dish and are ready to hold the CNT/PDMS composite (Gandhi and Raghava 2020; Jung et al. 2012; Liu et al. 2015). This technique is economical as several electrodes can be fabricated by re-using a single mold. Also, the time making the molds is not much time-consuming.

2.3 Composite formation: G-I, G-II, and G-III

2.3.1 Dispersion of nanomaterials with IPA

As discussed above, CNTs can form aggregates quickly. The clusters formed create a hindrance in the conductivity. To overcome this issue, CNTs should be properly dispersed to form an excellent conductive network. Dispersion is a technique in which the CNTs are mixed with IPA so that all the individual particles and molecules of CNT are separated in such a way it becomes one clear solution. The dispersion was achieved by bath-sonicator (PCI analytics) for 30 min, @50 Hz, and 230 V for solutions G-I, G-II, and G-III.

Thirteen solutions starting from 0.1wt%, 0.25wt%, 0.50wt%, 1wt%, 1.5wt%, 2wt%, 2.5wt%, 3wt%. 3.5wt%, 4w%. 4.5wt%, and 5wt% were prepared for all three groups, namely G-I, G-II, and G-III.

For preparing solutions for the G-I group, different concentrations of graphene as mentioned above were mixed with IPA and dispersed using a bath-sonicator for 30 min, at 50 Hz, and 230 V.

For preparing G-II solutions, Graphene and MWCNT-COOH were mixed according to the varied concentrations along with IPA and dispersed as discussed above.

Similarly, for preparing solutions for G-III which included Graphene and SWCNT-COOH, the two nanomaterials were mixed according to the required concentrations along with IPA and were sonicated using the same process as mentioned above.

2.3.2 Dispersion of nanomaterials with PDMS

To induce flexibility and give it a shape PDMS is mixed with the dispersed solutions. First, the PDMS base was dispersed with the above solutions with different concentrations for 10 min using the baht sonicator (50 Hz, 230 V). Next, the cross-linking agent was added and dispersed for 5 min with the same parameters. The cross-linking agent was added to cure the solution. Therefore, a total of thirty-nine electrode solutions were prepared with different concentrations including all three groups G-I, G-II, and G-III.

2.3.3 Fabrication of electrode

This is the concluding stage of the whole process. The CNT/PDMS composite is poured into the molds and kept in the oven for thermal curing. The temperature was set to 100 °C which was constant for all concentrations. Although, different concentrations are required to be in the oven for different time durations (2–10 h). After the solution was fully dried, the electrodes were separated from the molds (Fig. 2). Although most of the electrodes were separated easily, some of them required to be immersed in methanol before extraction. Lastly, for conduction purposes, steel snaps were affixed using the conductive Ten20 paste.

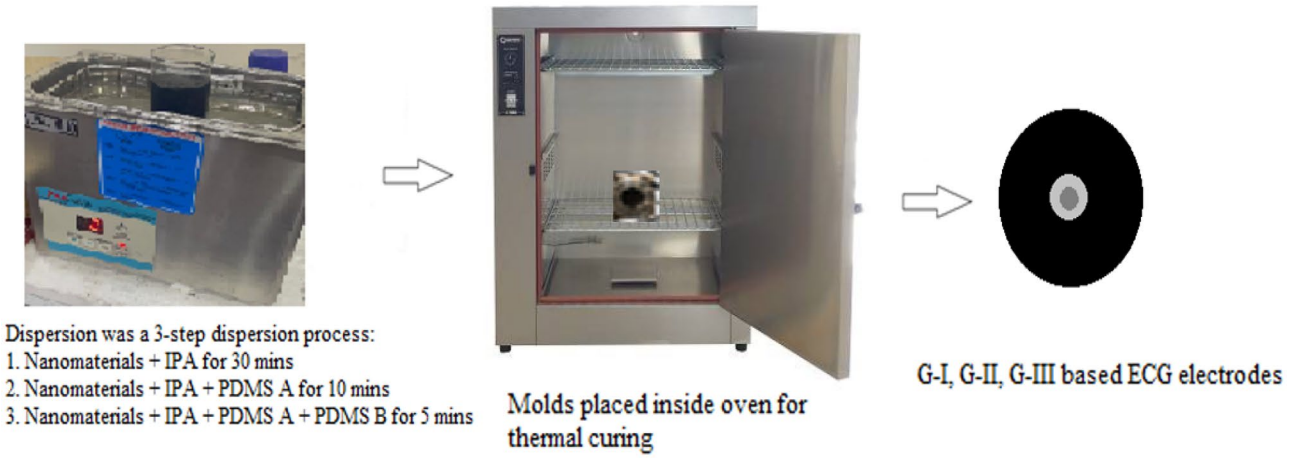


Fig. 2 Electrode fabrication process

3 Results

3.1 Physical characterization of MWCNT-COOH

MWCNT-COOH were characterized using Raman Spectroscopy, Fourier Transform Infrared Spectroscopy (FTIR), and Transmission Electron Microscopy (TEM).

Raman spectroscopy defines the structure as well as the molecular composition of MWCNT-COOH (Fig. 3).

The characteristic bands G-band and D-band are observed at 1334 cm^{-1} and 1557 cm^{-1} respectively. The G-band is allotted to the in-plane vibrations of the C–C bond. The G-band is the tangential shear mode of carbon atoms which links to the stretching mode in the graphite plane. The D-band at 1557 cm^{-1} is activated by the presence of some disorders in the C–C arrangements.

FITR was done to recognize the chemical bonds and covalent bonding signs between MWCNT and the -COOH group. The FITR was performed between 400 – 4000 cm^{-1}. The FITR graph (Fig. 2) offers data on the adsorption of waves by the -COOH functional group. The two peaks were at 1500 cm^{-1} and 3700 cm^{-1}, which means the bonds are stronger in this range, representing the presence of aromatic compounds.

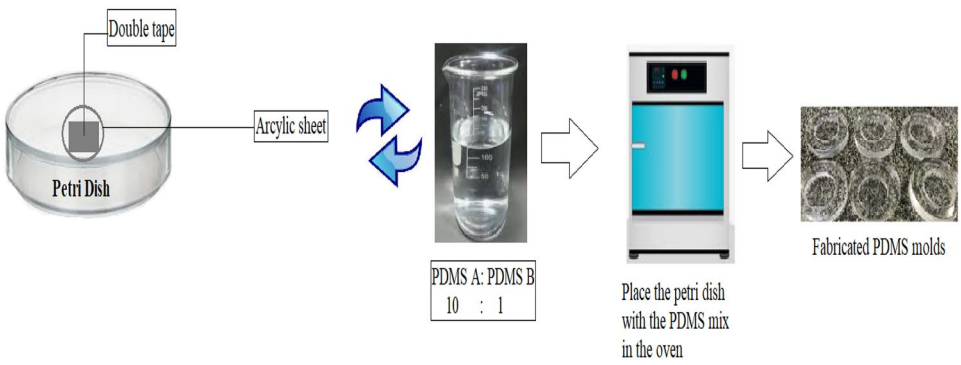
The subsequent topologic investigation was done using TEM physical characterization. TEM defines the internal structure of the composite. Here, the content of -COOH in MWCNT ranges between 2-4wt%. The topology suggests that the -COOH functional group breaks down the CNT bundles, which improves its solubility in Graphene as well as PDMS. As observed in Fig. 4, the MWCNTs are realized in bundles, but later the -COOH group is added to the CNTs isolated from each other, which increases the dispersion procedure and surface activities of the composite.

3.2 Physical characterization of SWCNT-COOH

SWCNT-COOH was also evaluated as a solid residue via several physical and chemical topological tests like Scanning Electron Microscopy (SEM), TEM, and Fourier Transform Infrared Spectroscopy (FTIR).

It is essential to characterize the composite to observe and analyze ruptures or adulterations on the surface. Henceforth, SEM was implemented for statistics on the composition, structure, and morphology of this composite. surface contaminations or fractures depicted in Fig. 5.

Fig. 3 Mold casting procedure



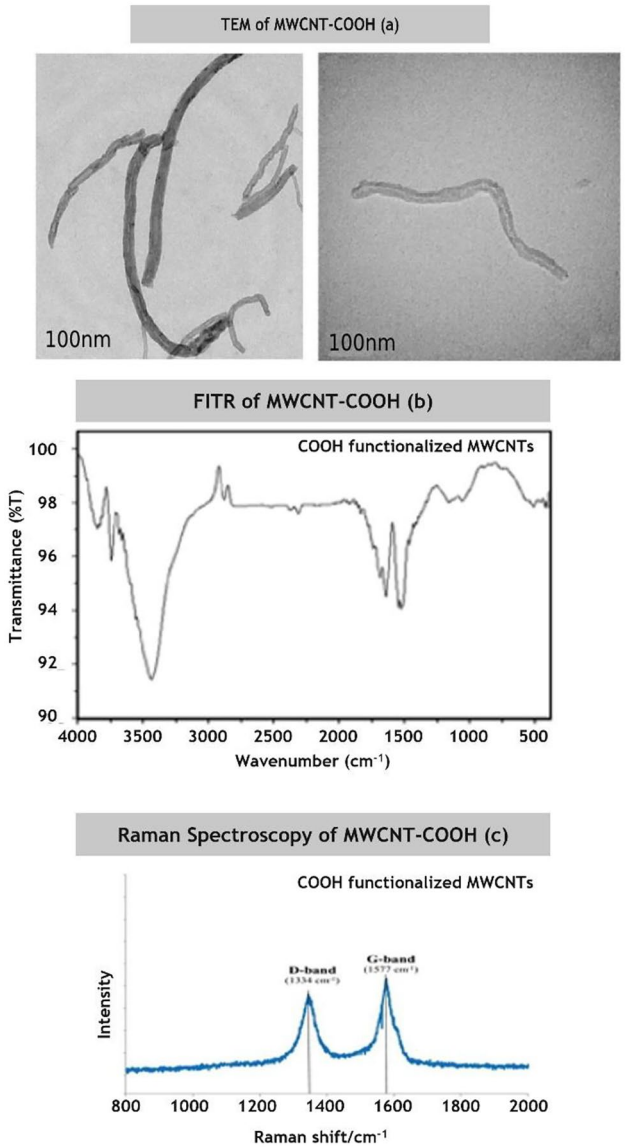


Fig. 4 Physical characterizations of MWCNT-COOH

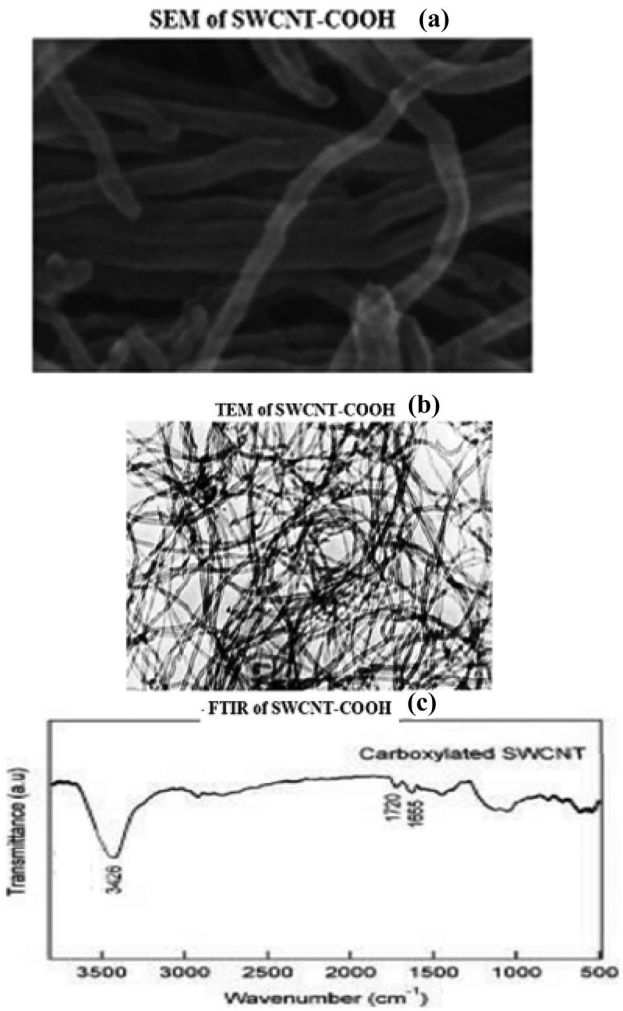


Fig. 5 Physical characterizations of SWCNT-COOH

While SEM offers evidence on the surface specifics and configuration, TEM offers statistics on the inner structures of the element. Hence, TEM was performed for SWCNT-COOH and is depicted in Fig. 3.

FTIR was performed to identify the chemical bonds and illustrate covalent bond information in this molecule by fabricating an infrared absorption spectrum. Three peaks were detected at 3425 cm-1, 1720 cm-1, and 1655 cm-1, which means the bonds are stronger in this range, representing the existence of aromatic compounds.

3.3 Smart ECG tracking framework

A generic architecture of an intelligent ECG tracking framework has been depicted in Fig. 1. It describes the principal elements and elementary modules of an intelligent ECG monitoring network. The system has been realized using, AD8232, NodeMCU, Arduino IDE, API, and AWS cloud services. Initially, the electrodes are connected to the volunteers on the right wrist, left wrist, and right ankle. The three wires of the cable are then connected to the data acquisition circuit (AD8232). Later, after the signal has been filtered and processed, it is sent to the IoT-based development board (NodeMCU). This board is equipped with a Wi-Fi module. In the next step, NodeMCU was programmed using the Arduino IDE, and the connection to the Wi-Fi was made with the login ID and password. Then the ECG signals were visible on the laptop screen. Then, the ECG value was pushed to the API for data transfer using 'HTTP Post'. Also, the latest value from NodeMCU is pushed to the API on the AWS cloud platform. Also, at the backend, the value is coming from the NodeMCU via 'HTTP Get'. After the value has been received, the application displays the ECG graph.

The ECG signal can also be viewed using the web browser. A database for storage of the data was also used, MongoDB.

3.4 Short-term monitoring

As the name suggests, the volunteer carried the electrodes for 5 h at the electrode sites. The first testing was done after immediately attaching the electrodes, and later after 5 h. The results were no different from each other. In addition to this, the skin-combability test was done. The electrodes were safe to wear, there were no skin rashes, irritations, or allergies seen on the skin site for all the three groups G-I, G-II, and G-III.

3.5 Long-term monitoring

Long-term ECG monitoring involved wearing the electrodes for 5 days in a row. The first piece of testing was done on the 1st day itself, followed by the 3rd day, and 5th day. The results were a bit degraded on the 3rd day, and 5th day as compared to the 1st day. Even after 5 days, the electrodes were proved to be skin-safe, as any sort of rashes or allergies were not observed for all the three groups G-I, G-II, and G-III.

3.6 Capacitance of the ECG electrodes: measurements and analysis

Capacitance is an indicator of the electron flow in the conductor, and it gives an idea about the charge. The capacitance was tested with the help of ESCORT ELC-133A. The capacitance depends on various factors such as dimensions of the conductor (electrode in this case), and material type. To evaluate this, a graph has been plotted between the concentration and capacitance of each electrode. As seen in Fig. 6, the capacitance is increasing for G-II (Graphene/MWCNT-COOH/PDMS), which is a good trend indicating that with increasing concentration, capacitance is increasing and the highest value of 0.061nF has been observed at 5wt%.

For G-III (Graphene/SWCNT-COOH/PDMS), the capacitance is increasing to 3 wt% on which the highest value was observed (0.172nF) and after that, the capacitance started to decrease stating that after a certain concentration of nanomaterials percolation threshold was achieved. This value was greater than the G-II value even at 5wt%.

However, for G-1 (Graphene/PDMS) the capacitance is almost the same it is almost throughout. It is almost a straight line but the values are increasing, although not significantly, which is again a positive sign. The highest value is observed at 5 wt% with 0.0176 nF.

The highest value of capacitance is observed with G-III at 3wt% with 0.172nF. The values achieved by G-I and G-II even at their highest concentrations are low as compared to the G-III group, this means that higher capacitance can be achieved at lower concentrations by using G-III elements.

3.7 Resistance of the ECG electrodes: measurements and analysis

Surface resistance is another important parameter as it indicates the opposition in the current flow of the electrode. The lesser the resistance, the better results are expected from the electrodes. The resistance was measured by completing the circuit, i.e., applying the voltage to the circuit. The probes of

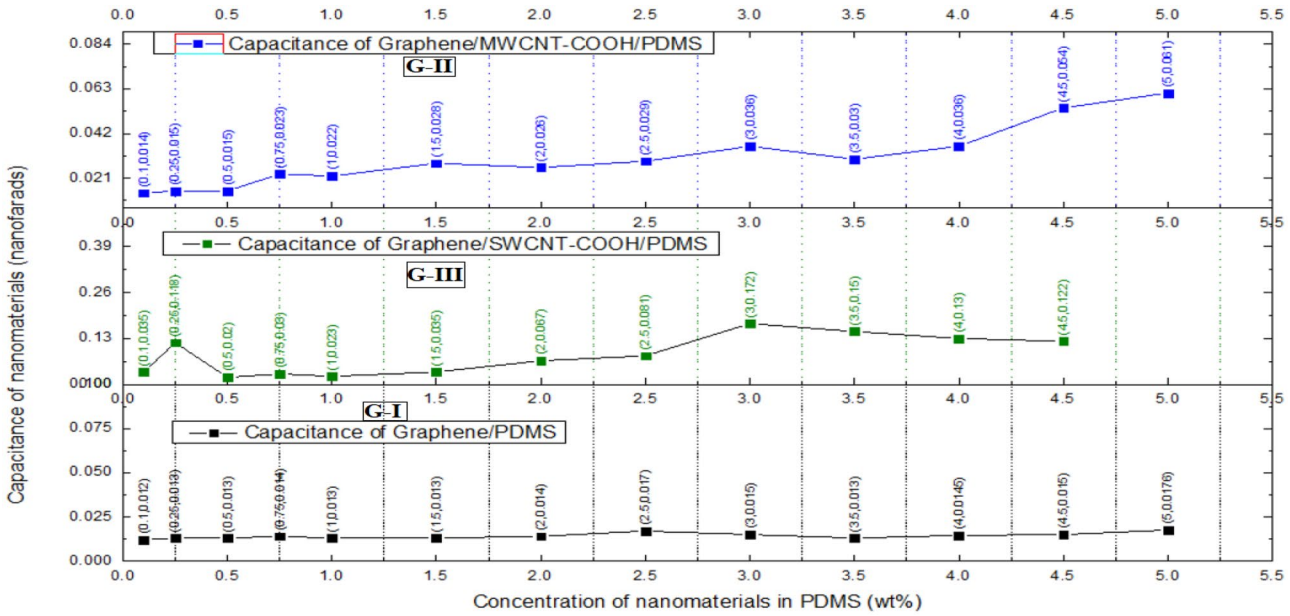


Fig. 6 Capacitance of G-I, G-II, and G-III

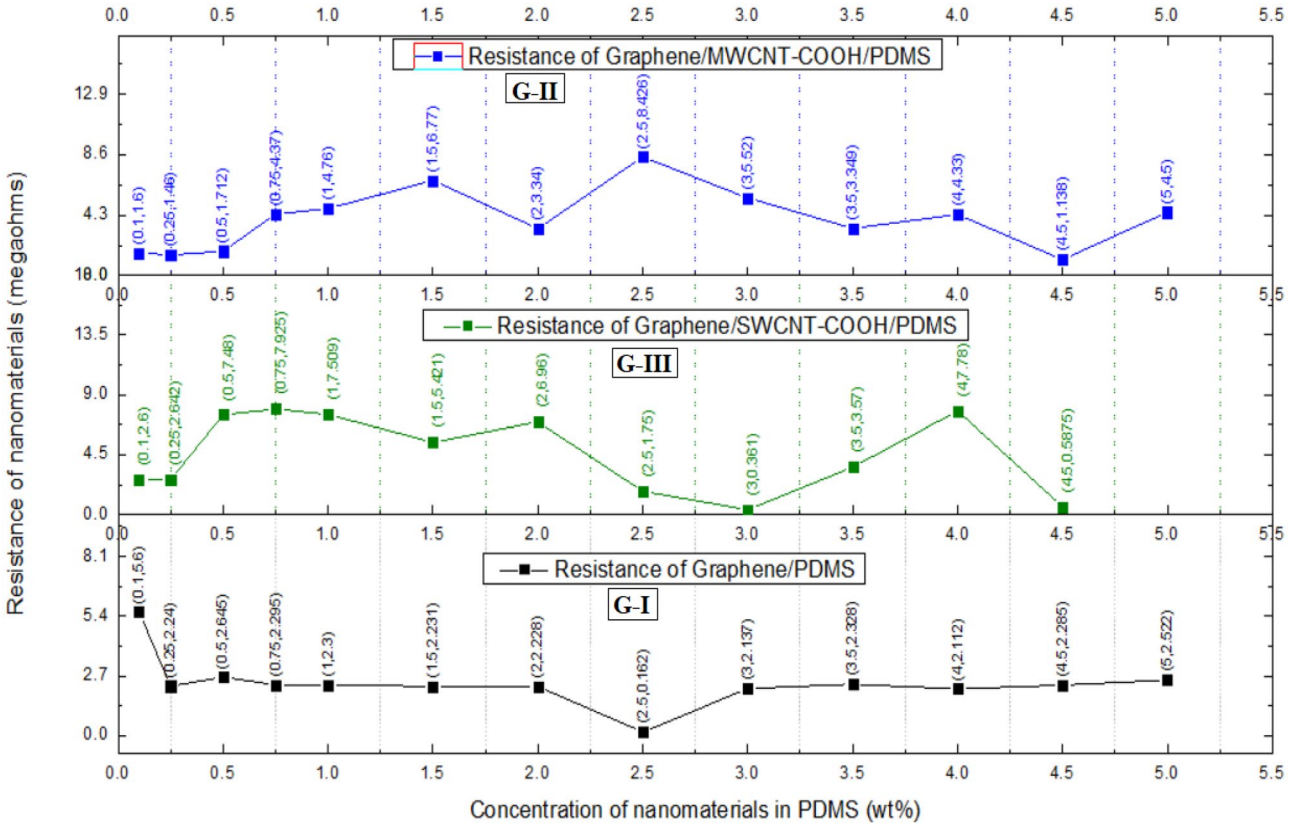


Fig. 7 Resistance of G-I, G-II, and G-III

the LCR meter (ESCORT ELC-133A), were connected to the two ends of the electrodes. A graph has been plotted between the surface resistance and various concentrations of electrodes.

As can be observed from the Fig. 7, for G-II, the resistance is decreasing as the concentration increases, but a peak resistance of 8.426 MΩ is observed at 2.5 wt%. The least

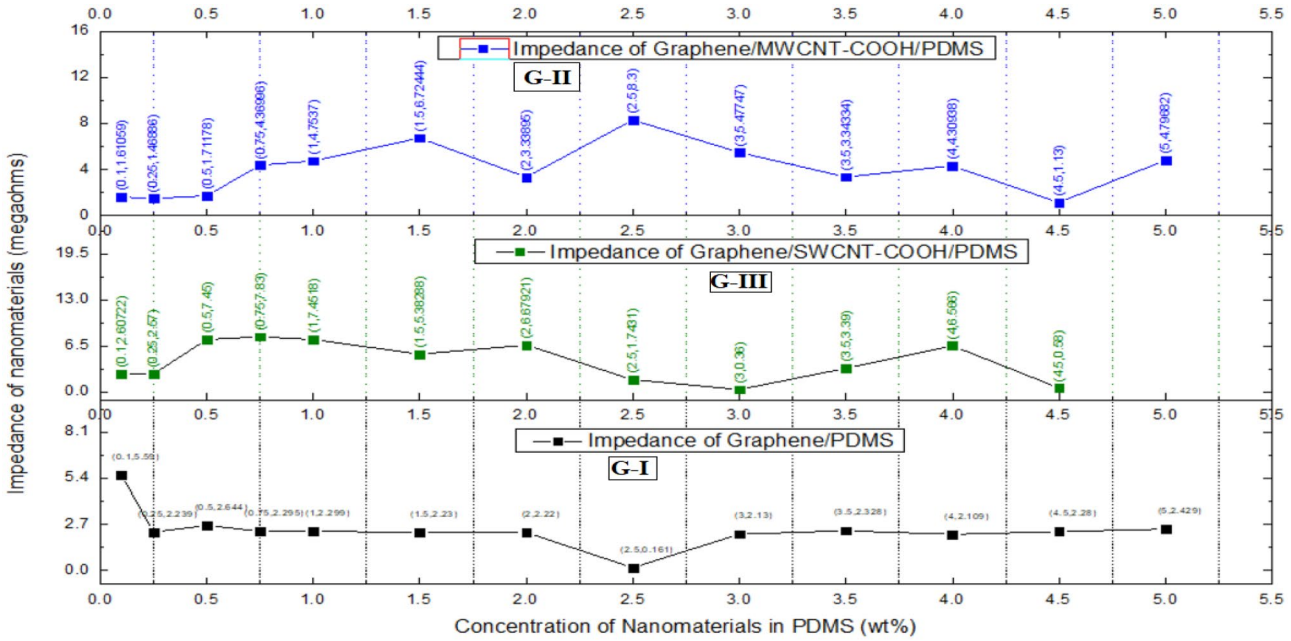


Fig. 8 Impedance of G-I, G-II, and G-III

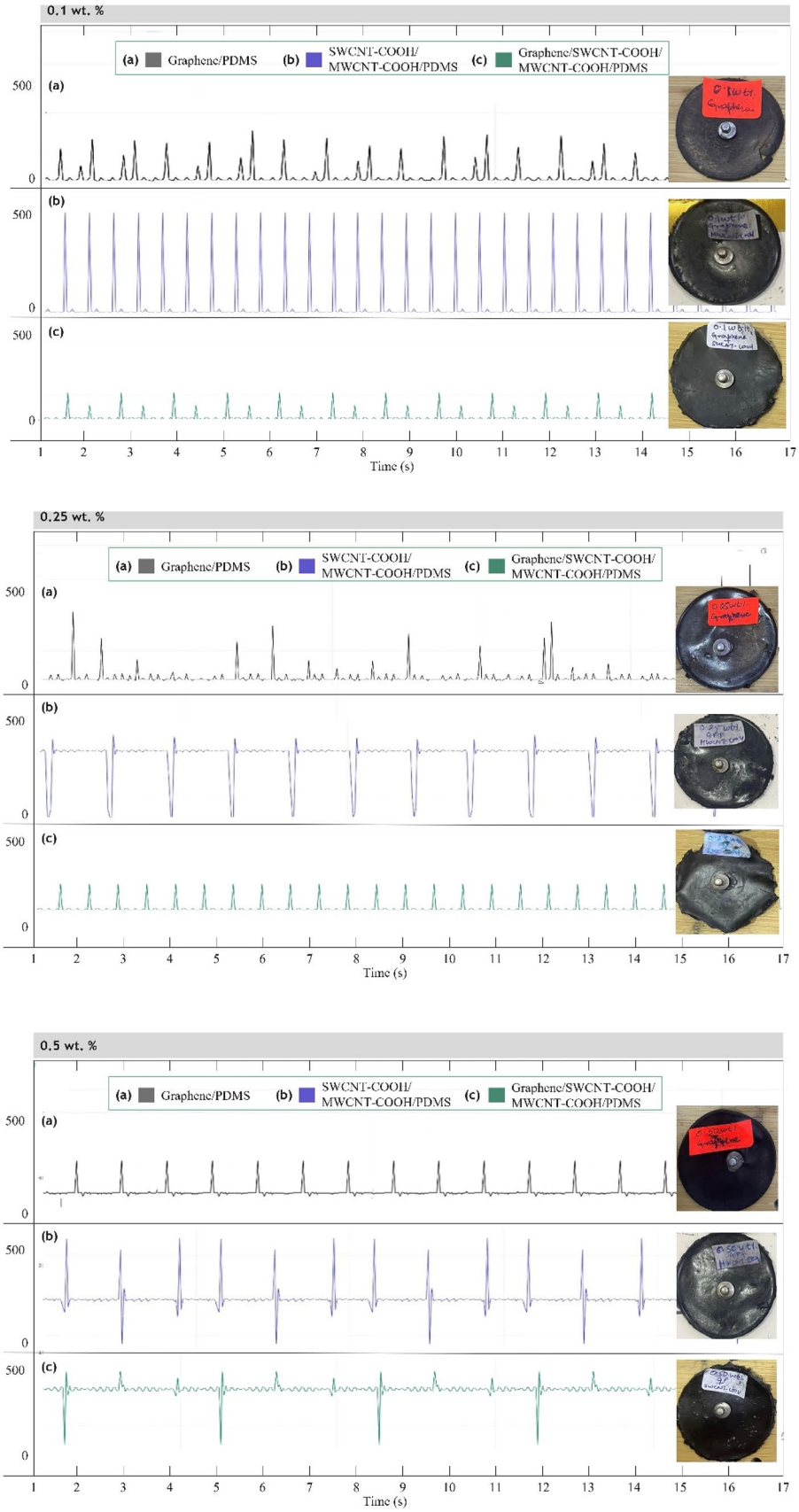


Fig. 9 Comparison of Different Electrodes

Fig. 9 (continued)

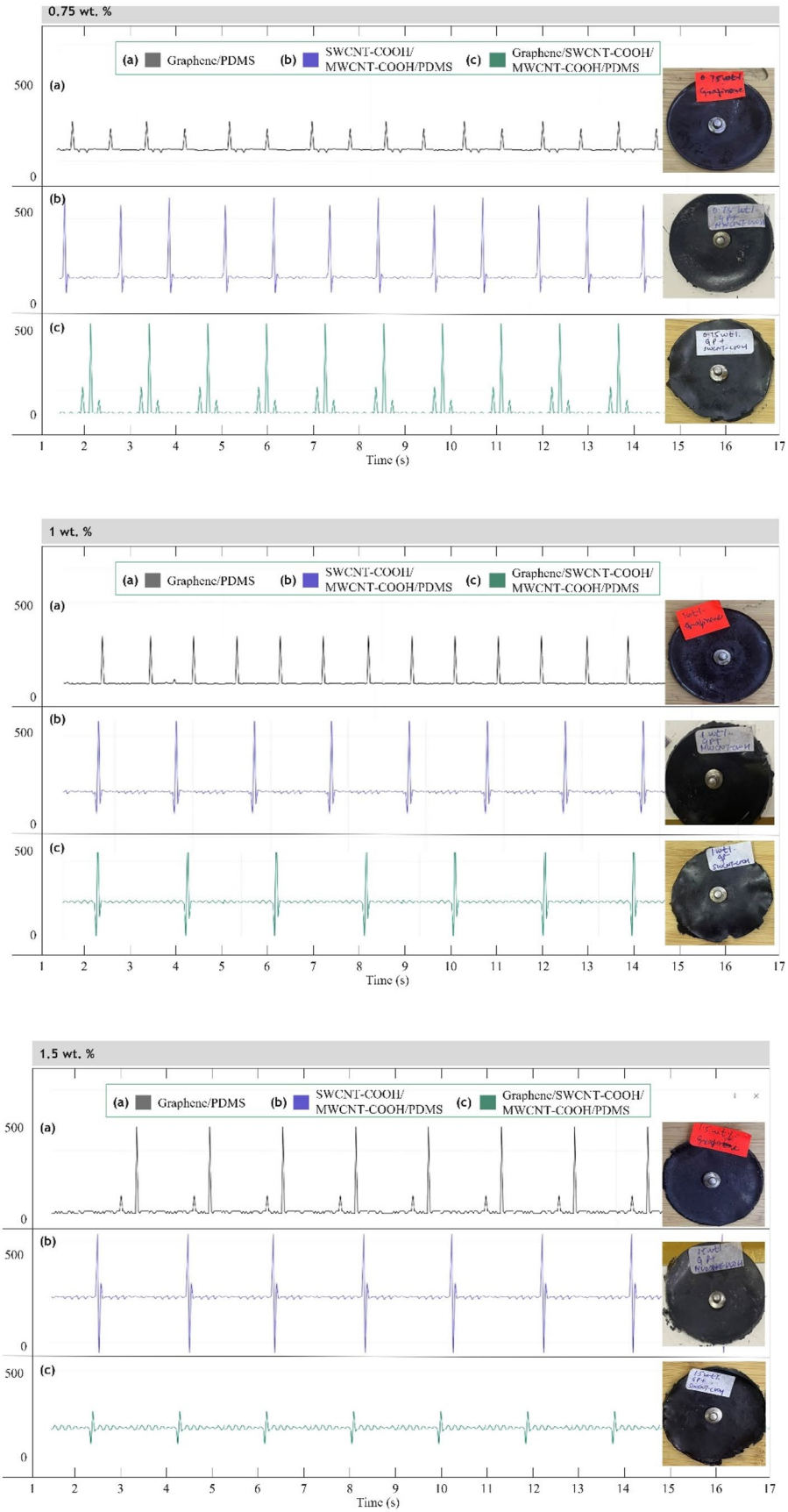


Fig. 9 (continued)

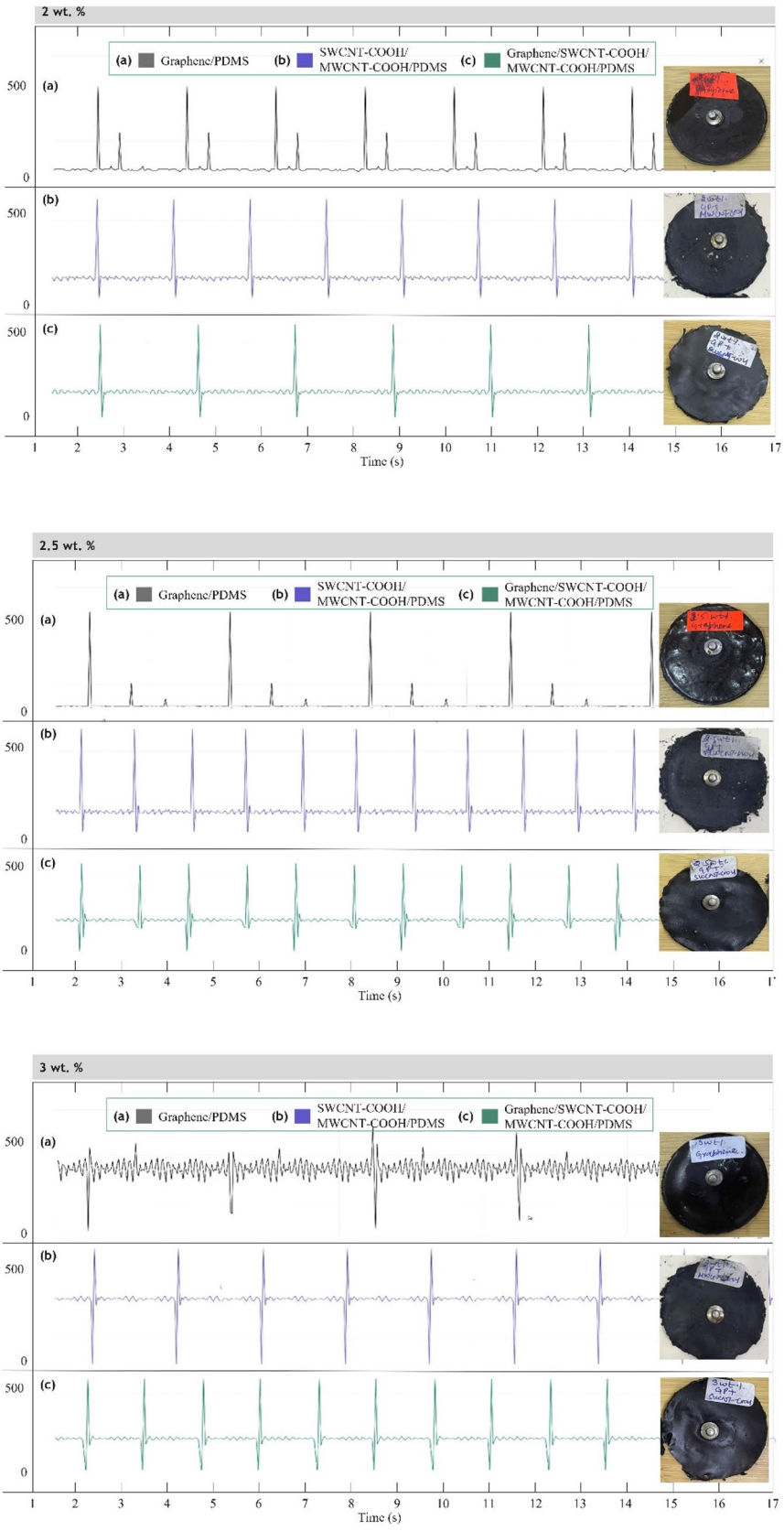


Fig. 9 (continued)

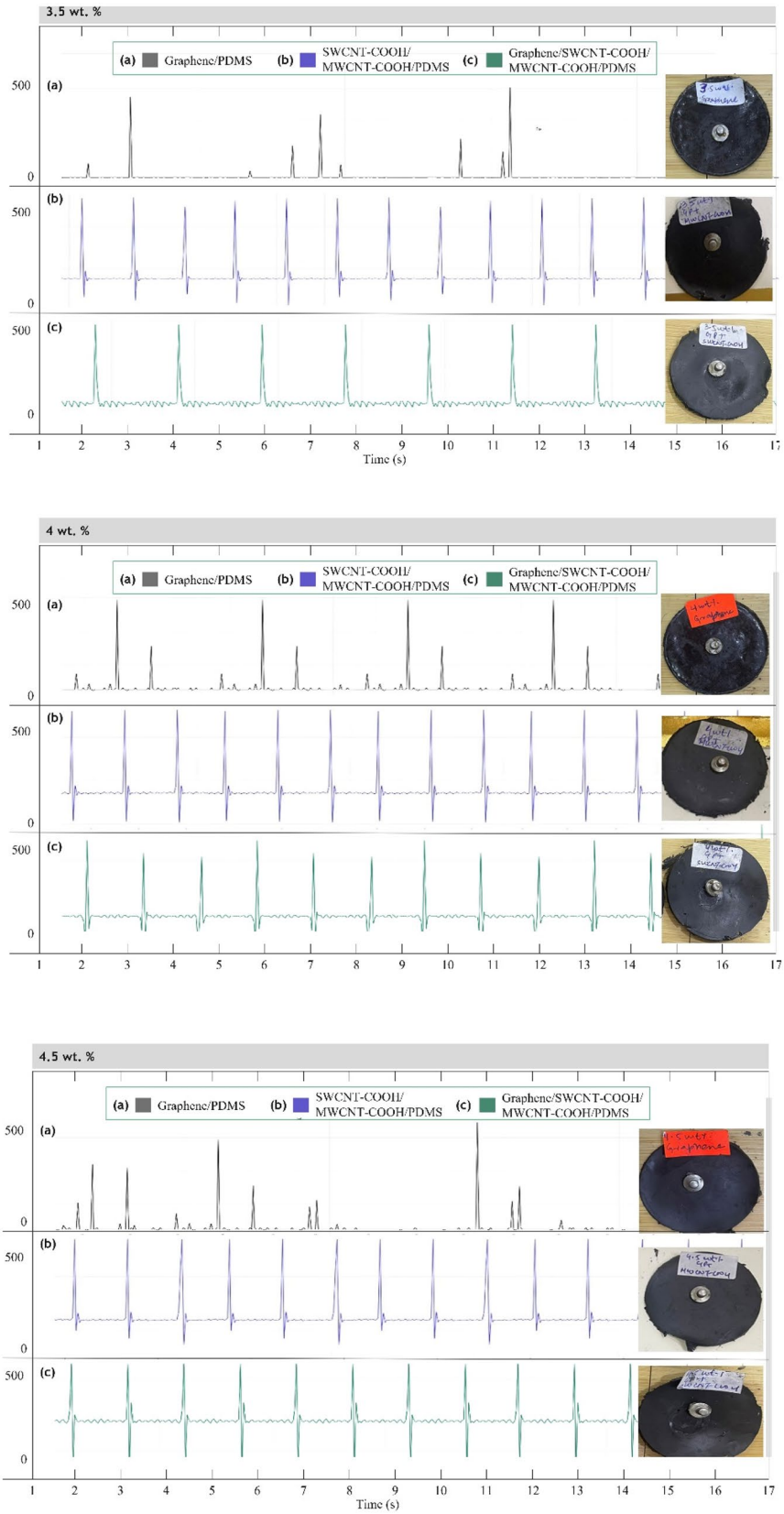
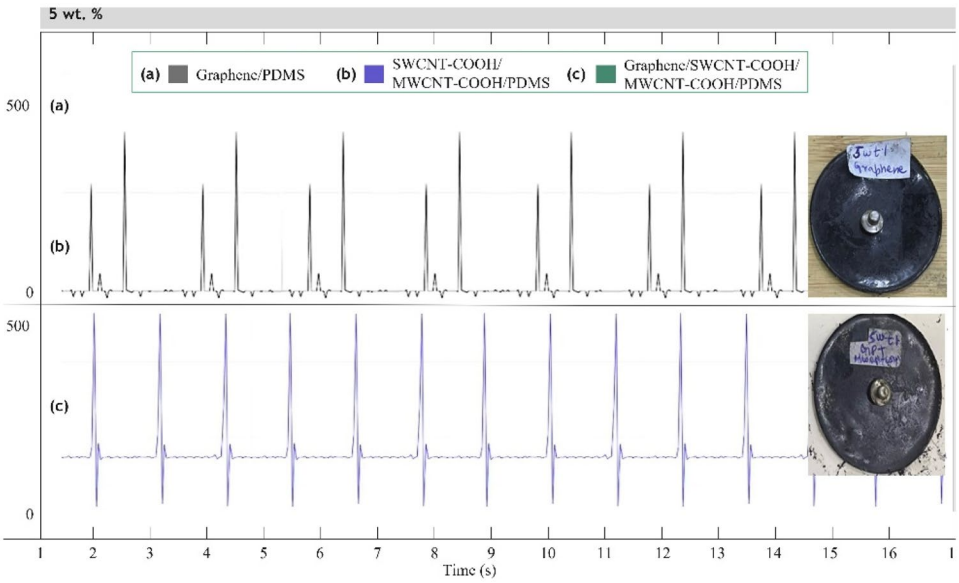
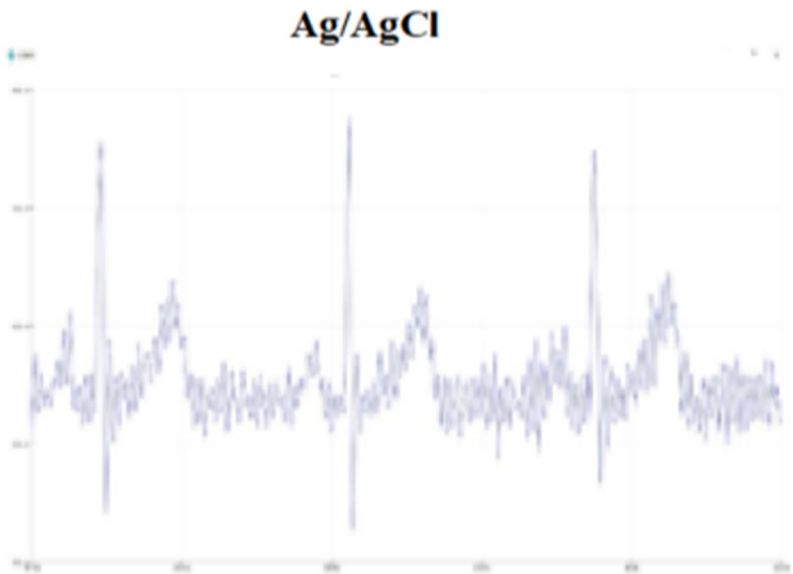


Fig. 9 (continued)





value of resistance 1.138 (MΩ) is observed at 4.5wt%, where the value of capacitance is 0.054nF, which shows an indirect relation between the two evaluations. Therefore, better results are achieved with higher concentrations.

For G-III the resistance is constantly decreasing, with an increase in nanomaterial concentration, but a sharp peak (7.78 MΩ) was observed at 4wt%, and the value of resistance dips again, which is a good trend.

For G-I, the values of resistance are not decreasing with increasing concentration, the values are almost constant at ~2 MΩ. This means the values of resistance are decreasing due to functionalized CNTs present in the composite.

3.8 Impedance of the ECG electrodes: measurements and analysis

As compared to wet electrodes, large amounts of capacitive elements are present in dry electrodes. Therefore, biosignals like ECG, impedance measurements should be taken into account so that signal distortion, signal attenuation, and motion artifacts can be taken care of. Therefore, to check the impedance concerning each of the samples fabricated, a graph has been plotted. The impedance was calculated at 100 Hz.

As it can be observed from the Fig. 8, for G-II, the impedance is decreasing as the concentration increases

Table 2 Comparative results of various works

S.No	Ref	Materials used	Fabrication technique	Nature of sensor	Comparative results	IoT-enabled
1.	Hossain et al. (2019)	Chemically-modified graphene/MWCNT-COOH	Drop cast	Film (circular)	CG/f@MWCNTs-25% portrayed best ECG results among others.	No
2.	Kim et al. (2016)	Graphene/MWCNT	Spin coating	Film (circular)	Results were comparable to Ag/AgCl electrodes.	No
3.	This work	G-I	Mold casting	Film (circular)	No definite ECG waveform was observed.	Yes
		G-II			2 wt% had better results than Ag/AgCl electrodes G-I and G-III.	
		G-III			3 wt% electrodes had results comparable to Ag/AgCl electrodes.	

but a peak of 8.426 MΩ is observed at 2.5 wt%. The least value of impedance (1.138 MΩ) is observed at 4.5wt%, where the value of capacitance is 0.054nF, which shows an indirect relation between the two evaluations. Therefore, better results are achieved with higher concentrations.

For G-III the impedance is constantly decreasing, with an increase in nanomaterial concentration, but a sharp peak was observed at 4wt% with 7.78 MΩ, and the value of impedance dips again, which is a good trend.

For G-I, the values of impedance are not decreasing with increasing concentration, the values are almost constant ~ at 2 MΩ. This means the values of impedance are decreasing due to functionalized CNTs present in the composite.

3.9 Comparison with Ag/AgCl electrodes

The ECG results were simulated from the system developed as discussed earlier. Figure 9, shows the results from various fabricated electrodes. The ECG results obtained from lower concentrations of the composite do not portray good ECG waveforms, i.e., the peaks are not visible properly.

For G-I the ECG waveform was not at all clear and was not even close to the conventional Ag/AgCl electrodes. The peaks were also not observed for any of the electrode concentrations.

For G-II the best ECG waveform was observed with clear peaks and equal amplitudes for all the intervals were seen at 4wt%.

For G-III the best ECG waveform was observed with clear peaks and equal amplitudes for all intervals were seen at 3wt%. Also, the percolation threshold was observed at 4.5 wt%, and no proper electrodes were fabricated for this group.

4 Discussion

In this section of the paper, a comparison table is presented with other works (Table 2). In Hossain et al. (2019) the authors used various equipments like sonicator, autoclave, oven, and a drop caster. The fabrication method required a longer duration and multiple equipments which are expensive and bulky. Also, the system is not IoT-enabled. In Kim et al. (2016) the mold was created using complex techniques like photolithography and etching. The fabrication was done using the spin-coting technique which is again an expensive equipment. Also, the system was not enabled for data transmission. In this work, the fabrication method is economical and user-friendly. No bulky and expensive equipments are required. Only a bath-sonicator and an oven is required for fabricating the electrodes and the molds. Also, the molds are casted using a simple procedure. The system is IoT-enabled and ECG obtained is of good quality.

5 Conclusion

Flexible ECG electrodes (G-I, G-II, and G-III) were fabricated with different concentrations from 0.1wt%-5wt%, but the percolation threshold was achieved at 4.5wt% for G-III electrodes. The ECG was monitored using an IoT-based tracking system. The results were compared with conventional ECG electrodes. For G-I electrodes the ECG results were not comparable to the conventional Ag/AgCl electrodes, also the values of capacitance did not show any significant improvements as the values were constant. For G-II electrodes, 4wt% were comparable with the conventional electrodes. The best ECG waveform with the sharpest peaks was seen 3wt% concentration of G-III ECG electrodes. With 3 wt%, concentration the ECG best results were received out of the three groups. Also, electrodes are fabricated with less concentration of nanomaterials which would help fabricate economical electrodes, with the best peaks and great conductivity it offers. Also, these electrodes offer least resistance, which again is an important parameter for acquiring the weak ECG signal. The electrodes (G-I, G-II, and G-III) were flexible as these could be used for long hours and many

days. The electrodes are safe for human skin, as no skin allergies or irritations were observed. Also, there was no discomfort while wearing the electrodes even for 5 days. The electrodes are suitable for long-term and remote ECG monitoring. Various tests conducted proved that at higher concentrations yield better results, but only up to a limit.

Declarations

Conflict of interests The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

J.Y. Baek, J.H. An, J.M. Choi, K.S. Park, S.H. Lee, Flexible polymeric dry electrodes for the long-term monitoring of ECG. Sens. Actuators. A. Phys. **143**, 423–429 (2008). https://doi.org/10.1016/j.sna.2007.11.019

B. Gandhi, N.S. Raghava, Fabrication Techniques for Carbon Nanotubes Based ECG Electrodes: A Review. IETE. J. Res. (2020). https://doi.org/10.1080/03772063.2020.1768909

M.F. Hossain, J.S. Heo, J. Nelson, I. Kim, Paper-Based Flexible Electrode Using Chemically-Modified Graphene and Functionalized Multiwalled Carbon Nanotube Composites for Electrophysiological Signal Sensing. Information. **10**(10), 325 (2019). https://doi.org/10.3390/info10100325

H.C. Jung, J.H. Moon, D.H. Baek, J.H. Lee, Y.Y. Choi, J.S. Hong, S.H. Lee, CNT/PDMS composite flexible dry electrodes for long-term ECG monitoring. IEEE. Trans. Biomed. Eng. **59**(5), 1472–1479 (2012). https://doi.org/10.1109/TBME.2012.2190288

J.H. Kim et al., Simple and cost-effective method of highly conductive and elastic carbon nanotube/ polydimethylsiloxane composite for wearable electronics. Sci. Rep. **8**(1375), 1–11 (2018). https://doi.org/10.1038/s41598-017-18209-w

T. Kim, J. Park, J. Sohn, D. Cho, S. Jeon, Bioinspired, Highly Stretchable, and Conductive Dry Adhesives Based on 1D–2D Hybrid Carbon Nanocomposites for All-in-One ECG Electrodes. ACS. Nano. **10**(4), 4770–4774 (2016). https://doi.org/10.1021/acsnano.6b01355

S.M. Lee, H.J. Byeon, J.H. Lee et al., Self-adhesive epidermal carbon nanotube electronics for tether-free long-term continuous recording of biosignals. Sci. Rep. (2015). https://doi.org/10.1038/srep06074

B. Liu, Y. Chen, Z. Luo, W. Zhang, Q. Tu, X. Jin, A novel method of fabricating carbon nanotubes polydimethylsiloxane composite electrodes for electrocardiography. Biomater. Sci. Polym. E. **26**(16), 1–7 (2015). https://doi.org/10.1080/09205063.2015.1082807

J.P. Lu, Elastic properties of carbon nanotubes and nanoropes. Phys. Rev. Lett. **79**(7), 1297–1300 (1997). https://doi.org/10.1103/PhysRevLett.79.1297

S. Mallakpour, S. Soltanian, Surface functionalization of carbon nanotubes: fabrication and applications. RSC. Adv. **6**, 109916–109935 (2016). https://doi.org/10.1039/C6RA24522F

A. Nag, S.C. Mukhopadhyaya, J. Kosel, Flexible carbon nanotube nanocomposite sensor for multiple physiological parameter monitoring. Sens. Actuators. A. Phys. **251**, 148–155 (2016). https://doi.org/10.1016/j.sna.2016.10.023

A. Pantelopoulos, N.G. Bourbakis, A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE. Trans. Man. Cybern. **40**, 1–12 (2010). https://doi.org/10.1109/TSMCC.2009.2032660

R.S. Ruoff, J. Tersoff, D.C. Lorents, S. Subramoney, B. Chan, Radial deformation of carbon nanotubes by van der Waals forces. Nature. **364**, 514 (1993). https://doi.org/10.1038/364514a0

The Dow Chemical Company, Technical Datasheet (2017). https://www.dow.com/content/dam/dcc/documents/en-us/productdatasheet/11/11-31/11-3184-sylgard-184-elastomer.pdf?iframe=true. Accessed 7 May 2017.

M.-F. Yu, O. Lourie, M.J. Dyer, K. Moloni, T.F. Kelly, R.S. Ruoff, Strength and Breaking Mechanism of Multiwalled Carbon Nanotubes Under Tensile Load. Science. **287**(5453), 637–640 (2000). https://doi.org/10.1126/science.287.5453.637

Z. Zhu, An Overview of Carbon Nanotubes and Graphene for Biosensing Applications. Nano. Micro. Lett. **9**(25), 1–24 (2017). https://doi.org/10.1007/s40820-017-0128-6

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835880

Grey wolf and Fuzzy logic based enhanced clustering approach for energy efficient operation of wireless sensor networks

Hardik Rohilla
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
hardikrohilla_2k18co137@dtu.ac.in

Himanshu
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
himanshu_2k18co148@dtu.ac.in

Harish Kumar
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
harishkumar_2k18co139@dtu.ac.in

Pawan Singh Mehra
Dept. of Computer Science & Engineering
Delhi Technological University
Delhi, India
pawansingh@dtu.ac.in

***Abstract*—A Wireless Sensor Network constitutes a collection of a large number of sensor nodes garrisoned in a given determined area so as to gather information and environmental data. They prove to be additionally beneficial in accumulating data from harsh environments where human intervention is scanty. The nodes are equipped with sensory, communication and computational capabilities which run on a battery which cannot be replaced effortlessly. Therefore, it becomes immensely important to work out an efficient, systematic and well-structured approach to inhibit power dissipation in the SNs so as to prolong the network lifetime, stability and service quality. This work puts forward a Grey Wolf Optimisation(GWO) based procedure for cluster head selection and then employs fuzzy logic(FL) for efficient cluster development. The fitness parameters used in cluster head selection are remnant energy, standard intra cluster stretch, separation from sink, balancing factor and node degree. On the other hand, fitness of the CH and separation between node and a CH are the two parameters which serve as crisp inputs into the fuzzy inference system(FIS) in order to decide the cluster formation mechanism. On simulation of the proposed protocol and then comparing the results with corresponding protocols, our approach shows enhancements in the network performance metrics such as FND(First node death), residual energy, throughput.**

Keywords*—Wireless sensor networks, Grey Wolf Optimisation, Fuzzy logic, Internet of Things

I. INTRODUCTION

Wireless Sensor Network (WSN) is an assemblage of small embedded devices, which are called sensors that are capable of communicating wirelessly[5]. The wireless connection helps in advancing IoT applications and unlocks new horizons of collaboration among the nodes in order to best suit the conditions and expectations. The data collected by each node is relayed to the central station either through single hop communication or multi-hop communication[4]. From the central station, the data can then reach the intended user or program over the internet, which can be used for purposes such as research, IoT analytics, prediction, recommendation, etc. However, one major challenge which persists is the limited battery power constraint which calls for innovative solutions in terms of developing an optimal communication and routing technique so as to utilize the limited available resources in the best possible manner and get the network working in a stable condition. This issue has got the attention of many researchers working towards implementing and deploying streamlined IoT models in different domains.

Various mathematical algorithms and optimisation models are applied to the network to achieve high performance[6]. Grey Wolf Optimisation[8] and Fuzzy Logic model[10] are two such techniques. GWO gives better convergence and stability results when compared to PSO and GA. GWO turns out to be particularly better in case of large scale sensor deployments. Fuzzy Logic also provides satisfying solutions without the need of complex mathematical modeling as it derives its consensus from fuzzy models which are based on linguistic variables.

The contributions of our proposed technique are as follows:

1. The proposed approach addresses the energy issue in WSNs and propounds a power balanced and structured protocol for enhancing network lifetime and stability.
2. Grey-Wolf Optimization is employed for electing the most qualified CHs. GWO inhibits numerous capabilities which are well taken advantage of in the work to streamline the job of CH selection.
3. Various crucial fitness parameters are used in the fitness function of GWO. These are residual energy, average intra cluster distance, separation from sink, priority factor and CH overhead cost.
4. FIS is designed to aid the decision of non-CH nodes while associating with a CH.
5. Along with proximity to CH, fitness of CH also serves as input to the FIS for reducing congestion and improving quality of service.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

6. The proposed protocol is simulated and weighed up with traditional as well as recent state-of-the-art protocols.

The remaining sections of the paper are as described. Section 2 details the related work carried out in the field. Section 3 describes the system network model. Section 4 explains the proposed methodology. Furthermore, in section 5, we talk about the simulation environment and analyze the resultant outcomes. Lastly, conclusion is presented in section 6.

II. RELATED WORK

An important objective which needs attention is the power consumption minimization in WSN. The prime approach to address this problem is somehow able to transmit the data more efficiently using different techniques. For balancing the overall energy in a wireless sensor network among control and normal nodes, base station etc, Clustering algorithm is the most reliable algorithm present right now for this task.In the other words, it can be considered as one of those algorithms where control servers are master and common nodes as slaves whose work is to help in proper transmission of data to CN. Traditional routing techniques like [1-3] tend to focus mainly on control nodes (CN) selection with unique formulation and parameters. However [8], [7], [14-16] are some of the recent researches which provide advanced protocols for energy efficient operations.

Heinzelman[1] was the first one to introduce, LEACH. This particular algorithm is non-statically operated, which was a novel idea at that time. This algorithm ignores some other parameters for selection of control nodes which ultimately leads to unbalanced distribution of CNs. Ignoring some parameters for control nodes selections is one of the drawbacks of this algorithm leading to unequal overall CNs distribution. However other parameters are not taken into account while selecting the control nodes and as a result leading to the uneven control nodes distribution. Thus, Heinzelman et al. proposed a centralized protocol, LEACH-C. However this approach led to a problem where the nodes closer to the sink drained energy more quickly as compared to nodes far from the sink. In other words, it led to the problem of hotspots due to poor network grouping.

Smaragdakis et al.[2] studied that LEACH does not make good use of the energy heterogeneity of the sensor nodes. Usually, a percentage of sensor nodes possess additional energy which can be a result of the initial setting or operation of the network. Therefore Smaragdakis et al. proposed SEP for clustering in an energy varied WSN. With this enhancement, they were able to properly utilize the additional energy and thus, were able to improve the network lifetime. Lin [13] came forward with an approach "HEED" that is based on clustering and includes extra energy as the parameter for control node selection. The fact that this approach was able to exploit the multiple transmission energy availability at the sensor level, is one of its main advantages. However, it is not guaranteed to select nodes more optimally in terms of energy optimization and extension of the network life.

PS Mehra[4] has put forward an approach, LBECR (a load-balancing efficient clustering and routing). This protocol is designed with the aim to have closest to equal energy dissipation so that the overall load in the network is balanced. In this research, three FIS are used for selecting the head of cluster, forming of a cluster and data routing. The research shows an improvement in load balancing, energy average, delaying the time of deaths of the sensor nodes and overall network lifetime with the help of novel FIS and a more load-balanced efficient routing.

To solve the issue of providing QoS in WSN, Jaiswal et al.[6] proposed a Grey Wolf Optimisation based protocol which is used for selecting the CH in the network addressing QoS based routing for improved Cluster Head(CH) – Base Station(BS) intercluster routing. Results were compared with existing protocols like LEACH, GECR, ESO-LEACH on residual energy, period of stability, and others like network lifetime, delay and throughput etc.. This study proved to be adequate for the designing of WSN in IoT. One of the limitations we found in this study was that it didn't talk about the likelihood of a node being selected as the CH.

Singh et al.[11] worked on the heterogeneity model of the WSN and proposed an energy efficient cluster head selection technique for increasing the network lifetime. It is found to contribute to decreasing the overall energy utilization in the network and thereby decreasing the net load over the cluster head. However, this paper does not perform the load balancing in the process of designing a better CH selection in addition to data aggregation procedure during data compilation.

III. SYSTEM MODEL

The system model for the study is described in following sub-sections: Network model, Energy model, overview of GWO and Fuzzy model.

A. Network Model

For the simulation of our research, the network properties and characteristics are described as follows. For simplicity, the area of interest is considered as a 2D plane in which initially all the SNs are assigned random positions. The size of the network area is taken as M=100x100 m^2, while the number of nodes initially deployed are n=100. With the help of a broadcast message in communication range R_c, a SN can sense the nodes present in the vicinity of it. Each of the nodes has a UniqueID associated with it. Initially, each node has a battery capacity of E_0 =1J.

The sink node is installed at the centre of the field at coordinates of (50,50). Fig. 1 represents an instance of a network with nodes spread randomly across the field with BS at center.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

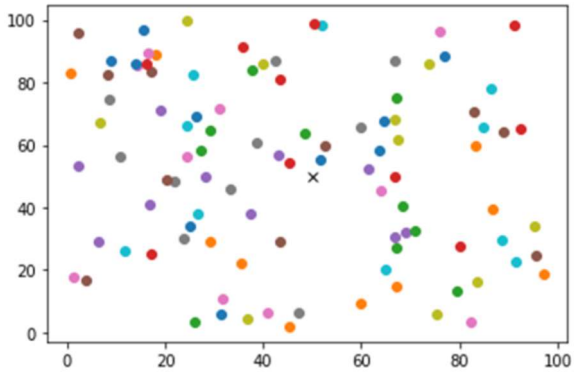


Fig.1 Random distribution of nodes with sink at center

B. Energy Model

The energy model adapted in our approach is that from Mehra et al.[4]. In order to determine a suitable energy model for estimating the performance of a technique, we need to take into account all the avenues around which a SN may be expending energy. In the operation of a round of communication, a SN can play either the role of a CH or a CM associated with a CH. CMs deplete their energy by transmission of sensed data to respective CHs using their radio electronics.[6] On the other hand, the job of CH is much more resource intensive as it depletes energy in the form of reception, data aggregation and transferral to BS. Therefore, the energy consumption in a CM for a round is given by the following set of eq 1.

$$E_{spent} = (E_{TX}*t) + (E_{fs}*t*D^2_{toCH}) \quad ,if\ D_{toCH} < d_0$$
$$E_{spent} = (E_{TX}*t) + (E_{mp}*t*D^4_{toCH}) \quad ,if\ D_{toCH} \geq d_0 \qquad (1)$$

Where E_{spent} is the energy exhausted in CM in transmission of a packet of length t over to a CH at a distance of D_{toCH}. E_{TX} is the energy exhaustion in electrical components for transmission. Efs and Emp represent the transmit amplifier types and accordingly the path loss factor i.e. 2 for free space and 4 for multipath fading.
The energy utilization in a CH node is given by the following set of equation 2.

$$E_{spent} = ((E_{RX}+E_{DA})*t*c) + (E_{TX}*t) + (E_{fs}*t*D^2_{toBS}) \quad ,if\ D_{toBS} < d_0$$
$$E_{spent} = ((E_{RX}+E_{DA})*t*c) + (E_{TX}*t) + (E_{mp}*t*D^4_{toBS}) \quad ,if\ D_{toBS} \geq d_0 \qquad (2)$$

Here, E_{RX} is the energy exhaustion in electrical components for reception. E_{DA} is the energy consumed for aggregation of 1 bit of information and c is the count of CMs associated with the CH. The threshold distance d_0 is defined as per equation 3.

$$d_0 = \sqrt{(Efs/Emp} \qquad (3)$$

C. Overview of GWO

Grey wolf optimization[8] is a pioneering metaheuristic algorithm developed by Seyedali Mirjalili et al. in 2014 with inspirations from leadership hierarchy as well as the technique of hunting by grey wolves in real life. The hierarchy (Fig. 2) in pack of grey wolves can be described as follow:

1. Alpha (α) – They are known for dominating other wolves of the pack and orders of them are obeyed by the pack members.
2. Beta (β) – They are one level below alpha or subordinate wolves known for helping the alpha wolf make necessary decisions for the pack. They are usually considered as the best candidates for becoming alpha.
3. Delta (δ) – They are suppressed by wolves on higher hierarchy but they suppress the omega wolves. There are different types of delta wolves like Hunters, Caretakers, Sentinels etc.
4. Omega (ω) – They are considered as the least important in the pack or scapegoat of the pack. They are dominated by every other category of the wolf and generally are allowed to eat at the end.

Phases of hunting:
1. Approaching
2. Encircling
3. Attack

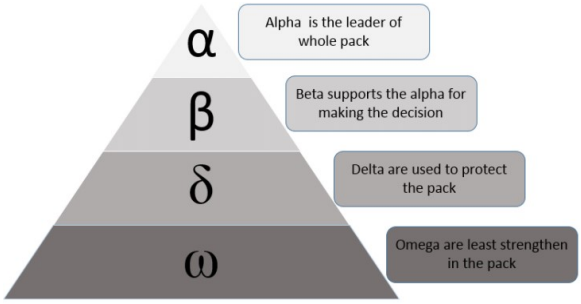


Fig.2 Hierarchy in a pack of grey wolves

The above grey wolf hunting behavior is mathematically modelled into Grey Wolf Optimization technique.[6] On the basis of how good or fit the solutions are, they are classified as different hierarchies. For example,

- Alpha wolf represents best solution as it may be in the proximity of a global best solution
- Beta wolf represents 2nd fittest solution
- Delta wolf represents 3rd fittest solution
- All other solutions are considered as the Omega wolves and they subsequently reposition them according to the positions of alpha, beta and delta wolves.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

Encircling of the prey:

The surrounding of prey can be described using following two equations 4 and 5:

$$\vec{D} = | \vec{C}.\overrightarrow{Xp}(t) - \overrightarrow{Xp}(t) | \quad (4)$$

$$\vec{X}(t+1) = \overrightarrow{Xp}(t) - \vec{A}.\vec{D} \quad (5)$$

here, t = current iteration

$\vec{A}$ and $\vec{C}$ = coefficient vectors

X_p = Position vector of Prey

X = Position vector of Grey Wolf

As the iterations are processed, $\vec{A} = 2ar_1 - a$ and $C=2r_2$ gets decreased progressively from 2→0 where r_1 and r_2 are arbitrary vectors within the range [0,1].

Hunting:

For hunting, omega wolves revise their positions according to their superior wolves, the reason being the preferable cognizance of the tentative location of the prey.[8] This is mathematically modelled using equations (6-12) summarised as below:

$$D_1 = |\overrightarrow{C1}.X(t) - \vec{X}(t)| \quad (6)$$

$$D_2 = |\overrightarrow{C2}.X(t) - \vec{X}(t)| \quad (7)$$

$$D_3 = | \overrightarrow{C3}.X(t) - \vec{X}(t) | \quad (8)$$

$$\overrightarrow{X1}(t+1) = X(t) - \overrightarrow{A1}.\vec{D} \quad (9)$$

$$\overrightarrow{X2}(t+1) = X(t) - \overrightarrow{A2}.\vec{D} \quad (10)$$

$$\overrightarrow{X3}(t+1) = X(t) - \overrightarrow{A3}.\vec{D} \quad (11)$$

$$\vec{X}(t+1) = (\overrightarrow{X1} + \overrightarrow{X2} + \overrightarrow{X3})/3 \quad (12)$$

Attacking:

To mathematically model the attacking behaviour, we use $\vec{a}$ by decreasing its value. $\vec{A}$ is nothing but an arbitrary value from [-2a,2a] where a gets linearly decreased from 2→0 for a number of iterations

|A|<1 – Attack the prey!

Searching:

|A|>1 – Retreat from the prey in order to locate a fitter prey.

Another factor of grey wolf optimization which can be considered for searching in C which has a random value within range of [0,2].

C >1 – Accentuate attack

C<1 – Deemphasize attack

D. Fuzzy Model

Whenever there is some precariousness involved in an exposition, FL[10] provides a pronounced way to model the system requirements and performance without a need of complex mathematical details. For the purpose of cluster formation, various factors like congestion, fitness, remaining energy and separation determine the appropriate CH for a non-CH node. So, by using FL we can map these parameters in simple linguistic terms and estimate the feasibility using IF-THEN rules.[4] These linguistic terms are formally known as fuzzy variables.[5] In the membership functions(MF), we have used trapezoidal MF for extreme variables and triangular MF for intermediary variables. These MFs range between 0 and 1.

Fig. 3 depicts the general flow of steps in a Fuzzy Inference System. Crisp input(s) are fed to the fuzzifier function, which returns the fuzzy variables for the input. This fuzzy set is then used to derive the output fuzzy set with the help of IF-THEN rules (framed according to the nature and requirements of the model). Finally, using the defuzzifier method (in our case, the CoS method[9]) we get the crisp numerical value associated with the output. The formula for the Center of Sum method is given by equation 13.

$$X* = (\sum_{i=1}^{n} x_i.A_{c_i}) / (\sum_{i=1}^{n} A_{c_i}) \quad (13)$$

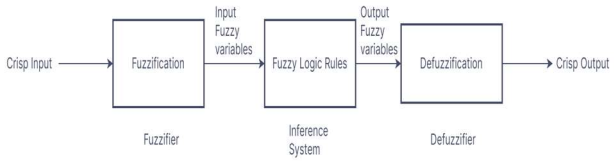


Fig.3 A generalised Fuzzy Logic based model

Here, A denotes the area of the region guided by a fuzzy variable and x denotes the central point of the area.

IV. PROPOSED APPROACH

The idea behind our proposed methodology is to make best use of the capabilities of GWO as well as FL to streamline the resource dissipation processes in such a way that the stability of the system can be increased. Two of the crucial processes in cluster based WSN architecture is the selection of CHs and formation of clusters. These two are discussed in the forthcoming section.

A. Selection of CHs

The algorithm 1 describes the steps involved in the CH selection mechanism using GWO. Each and every sensor node broadcasts a message so that it can be aware of the alive nodes present within its communication range. We iterate through all the alive nodes in the network one by one and run GWO in the communication range(R_c) of the node. Grey Wolfs are deployed within the area and after GWO completes its iteration, we have the fittest node in the area. Now, if the node in whose vicinity the GWO was run, is the most suitable node then it becomes a CH.

The percentage of CHs among alive nodes is set to p=10%. Therefore, we stop sooner when we have got the required number of CHs.

The fitness of a node is defined by a weighted function which is as defined in equation 14:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

$$F = (w_1 * f_1) + (w_2 * f_2) + (w_3 * f_3) + (w_4 * f_4) + (w_5 * f_5) \quad (14)$$

Here, w1, w2, w3, w4 and w5 are the weights for the parameters involved in the fitness function. These weights designate the priority of fitness components. The weights are equally distributed for all the fitness components which are normalised using min-max normalisation.

The parameters used in equation 14 are as follows:

1. Residual energy: The remnant energy of a node is a prime decider of its fitness to be a CH. Since, the role of CH consumes large amount of energy, therefore more the residual energy of a node higher is its chance of assuming the role of CH. In equation 15, $E_{residual}$ represents the current energy of the node.
$$f_1 : \frac{E_{residual}}{E_0} \quad (15)$$
2. Node degree: Larger the number of near neighbours, better is the feasibility of that node of being a CH. It is the ratio of the quantity of alive nodes in the sensing range and the amount of nodes in the cluster(N_c). In equation 16, V_i denotes the set of SNs within the sensing parameter..
$$f_2 : \frac{|V_i|}{N_c} \quad (16)$$
3. Average Intra cluster distance: Lesser the intra cluster distance, less is the communication cost overhead within a cluster and thus better is the clustering.
$$f_3 : \frac{\sum_{k=1}^{N_c} Dist(SN_k, SN_i)}{N_c} \quad (17)$$
4. Separation from the sink: The gap between BS and CH is also a significant parameter. Closer a node is to the BS, smaller is the transmission cost. (X_{BS}, Y_{BS}) and (X_{NODE}, Y_{NODE}) represent the coordinates of BS and the node in consideration respectively.
$$f_4 : \sqrt{(X_{BS} - X_{NODE})^2 + (Y_{BS} - Y_{NODE})^2} \quad (18)$$
5. Balancing factor: This factor ensures that the CHs are rotated suitably. The priority of a SN to be selected as a CH is inversely proportional to T_{CH} (the number of times it has assumed the role of CH previously).
$$f_5 : \frac{1}{1+T_{CH}} \quad (19)$$

Algorithm 1 : For suitable cluster heads selection

INPUT : Alive Sensor nodes in an area
BEGIN :
1 : LIST_NODES <- All alive sensor nodes
2 : SN(i) <- Sensor node with i as ID
3 : Every SN(i) broadcasts hello message in Rc
4 : For each SN(i) in LIST_NODES do
5 : Compute fittest node in range Rc
6 : If CH_Count<p% then
7 : If SN(i)==fittest then
8 : SN(i) becomes a CH
9 : CH_Count++
10: End If
11: End If
12: End For
TERMINATE
OUTPUT : Set of CHs selected for a round

B. Formation of clusters

Algorithm 2 describes the procedure for cluster formation. Cluster formation needs to take place when all the cluster heads are selected . According to the existing protocols for the cluster formation , a non cluster node has to make a greedy decision by joining with the closest Cluster Head Node. Now in order to help the decision of Non Cluster head nodes , An FIS (Fig. 4) that we designed which takes two inputs in order to achieve the chance. The two inputs that we are talking about are the cluster head's distance and suitableness for the cluster head for consideration . Since our goal is to minimize the energy loss of non - cluster head nodes that's why we consider distance and with the suitableness one can actually estimate the burden of Cluster Head node can take . Nodes which are non cluster heads received the telecast of selected Cluster heads announcing their campaigning during the cluster head selection process.

Now non cluster head nodes will have the distance and sustainability of cluster heads in their region with the help of that telecast. Non cluster head nodes have to make a decision to join the preferable cluster head,in order to maintain the balancing of the load of the network. To support this decision the FIS that we designed is used.

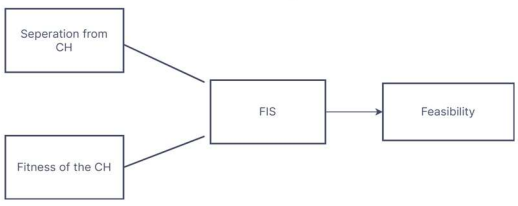


Fig.4 Designed FIS for cluster formation

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

Algorithm 2 For cluster formation mechanism

INPUT : Non clustered network
BEGIN :
1 : LIST_NODES <- All alive sensor nodes
2 : CH_NODES <- Sensor nodes chosen as CHs
3 : SN(i) <- Sensor node with i as ID
4 : For each non-CH node in LIST_NODES do
5 : For each SN(i) in CH_NODES do
6 : Compute feasibility <- fuzzy(seperation to CH,fitness)
7 : Send join ACK to CH with highest feasibility for establishment of connection
8 : End For
9 : End For
TERMINATE
OUTPUT : Clustered Network

Illustration of Fuzzy Rules:

Table 1 summarises the Fuzzy IF-THEN rules for all the possible permutations for the linguistic variables through which we can determine the fuzzy set for feasibility.

TABLE I Rules for Computing feasibility of a CH in order to select the best possible choice

Rule	Separation from CH	Fitness	Feasibility
1.	AtHand	Needy	Decent(D)
2.	AtHand	Satisfactory	Powerful(P)
3.	AtHand	Preferable	Massive Powerful(MP)
4.	Reasonable	Needy	Decent Low(DL)
5.	Reasonable	Satisfactory	Decent Powerful(DP)
6.	Reasonable	Preferable	Quite Powerful(QP)
7.	Extent	Needy	Massive Low(ML)
8.	Extent	Satisfactory	Quite Low(QL)
9.	Extent	Preferable	Low

V. SIMULATION AND RESULT ANALYSIS

The ensuing segment discusses the simulation environment used for implementation of the protocols and then, comparative analysis of results is presented.

A. Simulation Experiment

The experimentation for LEACH[1], GWBOC[6] and our proposed protocol has been carried out in python. Python is chosen because of its wide range of capabilities like portability, library support, data visualisation, etc. Overall, python serves as a good alternative to MATLAB when it comes to simulation of WSN backed IoT networks.

Sensor nodes are assigned fixed coordinates in a 2D plane spanning over the field from which information has to be gathered. One round of transmission is said to be completed when all the SNs which are capable of sending data, relays it to the centrally located base station. Moreover, a SN is said to be dead when it exhausts all of its initial energy as a result of participation in antecedent rounds. The simulations are carried simultaneously for all the protocols(i.e. for a single specific instance of node distribution in the network) to prevent any misleading results. Also, multiple sets of simulations are carried out to ensure reliability of the results). The various simulation parameters are shown via table II.

TABLE II Network parameters

Parameter	Value
Number of nodes(n)	100
Network area(M)	100 x 100
Sink coordinates ($sink_x$, $sink_y$)	(50, 50)
CH selection percentage	10%
Radio electronic energy	50 nJ/bit
Data Aggregation energy (E_{DA})	5 nJ/bit
Initial energy of a node (E_0)	1 J
Energy Heterogeneity level	1-level
Amplifier energy for multipath fading (E_{mp})	0.0013 pJ/bit/m^4
Data Packet size (t)	4000 bits
Maximum rounds (r_{max})	3000

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

B. Results analysis

We have simulated our proposed model against LEACH and GWBOC. The primary reason being that LEACH forms the base of all clustering protocols and it was one of the first basic algorithms which aimed to achieve energy optimisation in WSN. Since in GWBOC, there was no information of likelihood of a chosen node being the CH, we used GWO and our novel FIS to solve the issue and compared our results to examine how our model performed as compared to GWBOC. The following performance evaluation metrics are used for comparative study of different protocols with our proposed technique.

FND, HND and LND

The indication of a good protocol can be interpreted from its performance based on FND, HND and LND. These give the overall picture of the efficiency of an approach in utilising SN energy for prolonging its operation. Table III details the node lifetime statistics. FND of our proposed protocol depicts improvement from LEACH by 34.22% and from GWBOC by 24.89%. HND of our proposed protocol depicts improvement from LEACH by 48.95% and from GWBOC by 1.37%. LND of our proposed protocol depicts improvement from LEACH by 52.64% and from GWBOC by 14.15%.

TABLE III Network lifetime comparison in terms of FND,HND and LND

Approach	FND	HND	LND
LEACH	1499	1681	1928
GWBOC	1611	2470	2578
Proposed	2012	2504	2943

Fig. 5 shows the progression of rounds with the number of dead nodes in the network. We can see that LEACH reaches its last node death even before the FND occurs for our proposed protocol.

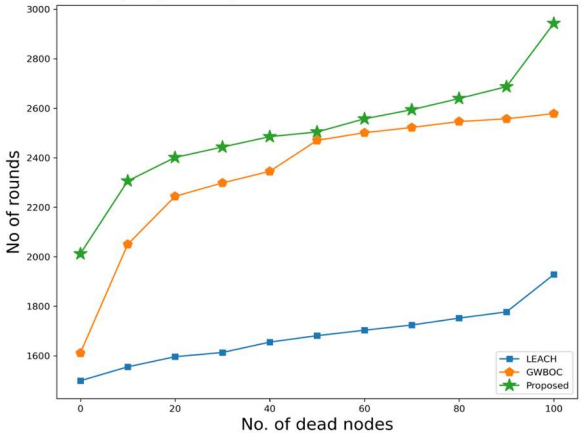


Fig.5 Comparison of the number of dead nodes with the progression of rounds

Throughput

Throughput of a network indicates the number of data packets administered to the sink or BS. When one packet of information from a CM node makes its way to the BS through a CH, the count of total data packets to BS is incremented by one. Through this procedure, we get the total throughput for all the three simulated approaches. Fig. 6 shows a graphical representation of the total number of packets sent to the BS as a consequence of the operation of the network. Our proposed protocol outperforms LEACH by 51.23% and GWBOC by 6.62%. Therefore, along with efficient energy dissipation in the SNs, our proposed protocol is able to send an adequate amount of sensory information to the BS.

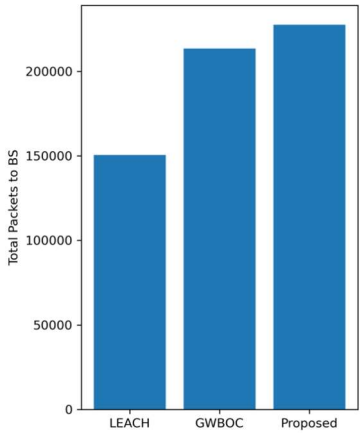
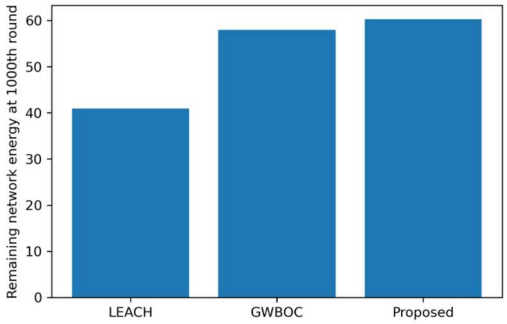


Fig.6 Comparison based on total number of packets sent to BS

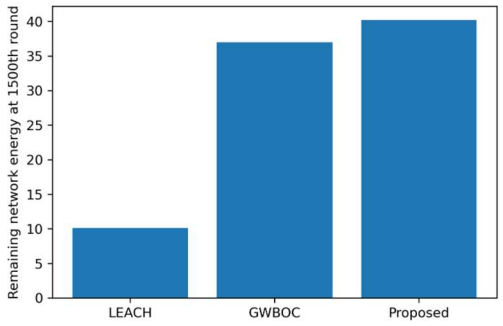
Network energy

In this section, we analyse the remaining energy of the nodes in the network after 1000, 1500 and 2000 rounds of functioning (Fig. 7(a)(b)(c)). Through this metric, we can clearly see that energy aware intelligent clustering leads to better power dissipation rate in the SNs.

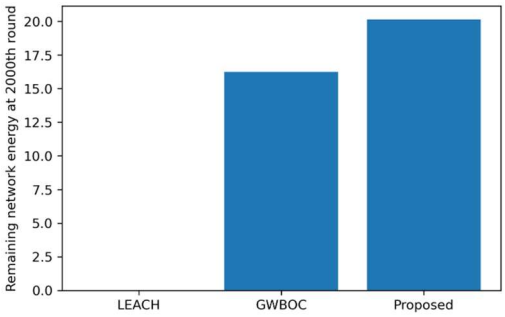


(a)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.



(b)



(c)

Fig.7 Remaining network energy comparison at various stages

VI. CONCLUSION

With the use cases of IoT getting increased in number as well as sophistication, the technology has found its way into our day to day lives and has been embedded deep in our routines, be it our professional work or recreational activities. So, in order to make complete use of the applications, the main issue is of best utilisation of resources and prolonging the lifetime. Researchers from all over the world are working relentlessly to confront this problem. In this study, we proposed a viewpoint which used GWO for efficient CH selection in the network and FL for cluster formation in the network. The technique is developed while keeping in consideration all the advanced measures of the eligibility of nodes to assume a particular role for a particular round based on the conditions within the network. The simulation is carried out extensively for our proposed approach and two prominent protocols namely, LEACH and GWBOC. The results reveal significant furtherance in network lifetime, throughput and network energy. As part of future work, we will work with the Fuzzy type-2 model[12] to further improve our findings. We are also aiming to work with mobile Sensor nodes network as this special type of deployment has a good scope and increased number of use cases at the industry level.

REFERENCES

[1] Heinzelman, Wendi B., Anantha P. Chandrakasan, and Hari Balakrishnan. "An application-specific protocol architecture for wireless microsensor networks." *IEEE Transactions on wireless communications* 1.4 (2002): 660-670.

[2] Smaragdakis, Georgios, Ibrahim Matta, and Azer Bestavros. "SEP: A stable election protocol for clustered heterogeneous wireless sensor networks." *Second international workshop on sensor and actor network protocols and applications (SANPA 2004)*. Vol. 3. 2004.

[3] Hong, Jiman, et al. "T-LEACH: The method of threshold-based cluster head replacement for wireless sensor networks." *Information Systems Frontiers* 11.5 (2009): 513-521.

[4] Mehra, Pawan Singh. "LBECR: load balanced, efficient clustering and routing protocol for sustainable internet of things in smart cities." *Journal of Ambient Intelligence and Humanized Computing* (2022): 1-23.

[5] Mehra, Pawan Singh. "E-FUCA: enhancement in fuzzy unequal clustering and routing for sustainable wireless sensor network." *Complex & Intelligent Systems* 8.1 (2022): 393-412.

[6] Jaiswal, Kavita, and Veena Anand. "A Grey-Wolf based Optimized Clustering approach to improve QoS in wireless sensor networks for IoT applications." *Peer-to-Peer Networking and Applications* 14.4 (2021): 1943-1962.

[7] Chithaluru, Premkumar, et al. "I-AREOR: An energy-balanced clustering protocol for implementing green IoT in smart cities." *Sustainable cities and society* 61 (2020): 102254.

[8] Mirjalili, Seyedali, Seyed Mohammad Mirjalili, and Andrew Lewis. "Grey wolf optimizer." *Advances in engineering software* 69 (2014): 46-61.

[9] Wang, Wen-June, and Leh Luoh. "Simple computation for the defuzzifications of center of sum and center of gravity." *Journal of Intelligent & Fuzzy Systems* 9.1, 2 (2000): 53-59.

[10] Zadeh, Lotfi A. "Fuzzy logic." *Computer* 21.4 (1988): 83-93.

[11] Singh, Samayveer, Aruna Malik, and Pradeep Kumar Singh. "A threshold-based energy efficient military surveillance system using heterogeneous wireless sensor networks." *Soft Computing* (2021): 1-14.

[12] Karnik, Nilesh Naval, Jerry M. Mendel, and Qilian Liang. "Type-2 fuzzy logic systems." *IEEE transactions on Fuzzy Systems* 7.6 (1999): 643-658.

[13] C.H. Lin, M.J. Tsai, A comment on 'HEED: A Hybrid, Energy-Efficient, Distributed lustering approach for ad hoc sensor networks, IEEE Trans. Mob. Comput. 5 10) (2006) 1471–1472.

[14] Jacob, I. Jeena, and P. Ebby Darney. "Artificial Bee Colony Optimization Algorithm for Enhancing Routing in Wireless Networks." Journal of Artificial Intelligence 3, no. 01 (2021): 62-71.

[15] Chen, Joy Iong Zong, and S. Smys. "Optimized Dynamic Routing in Multimedia Vehicular Networks." Journal of Information Technology 2, no. 03 (2020): 174-182.

[16] Ram, G. Mohan, and E. Ilavarsan. "Review on Energy-Efficient Routing Protocols in WSN." Computer Networks, Big Data and IoT (2021): 851-871.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:05:57 UTC from IEEE Xplore. Restrictions apply.

Environ Monit Assess (2022) 194:673
https://doi.org/10.1007/s10661-022-10340-0

Groundwater quality assessment and modeling utilizing water quality index and GIS in Kabul Basin, Afghanistan

S. K. Singh · Ali Reza Noori

Received: 11 December 2021 / Accepted: 8 August 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract Groundwater stands as a unique source of water supply in Kabul city, Afghanistan. In this investigation, 35 samples of groundwater were comprehensively analyzed to determine its hydrogeochemical characterizations, quality, water types, and its acceptability as drinking sources. A portable digital multiparameter instrument (LAB MAN Scientific instrument) was used to measure the total dissolved solids (TDS), hydrogen potential (pH), and electrical conductivity (EC). Total hardness, chloride, and bicarbonate were examined via a titrimetric approach. Sodium, calcium, magnesium, and potassium concentrations were measured with a flame photometer. Fluoride was determined by using a digital portable multiparameter. UV–VIS spectrophotometers were employed to count sulfate and nitrate concentrations. The distribution pattern of measured parameters and the Water Quality Index (WQI) in groundwater were spatially modeled utilizing the ArcGIS tool. The findings provide insight into the main anions and cations, which are found in ascending sequence $F < NO_3 < SO_4 < Cl < HCO_3$ and $K < Ca < Na < Mg$, respectively. Based on the measurements of ion concentrations, bicarbonate (71.4%), chloride (14.28%), nitrate (2.85%), magnesium (80%), sodium (82.85%), calcium (5.71%), and potassium (17.14%) were all determined to be over the World Health Organization (WHO) limits of drinking water. Using the Piper trilinear diagram, two significant hydrochemical facies ($CaNaHCO_3$ and $NaHCO_3$) were discovered. Based on the mathematical model of WQI outputs, 88.57% of the research region has excellent to good water, whereas 11.43% has poor to very poor water.

Keywords WQI · Groundwater quality assessment · Spatial modeling · ArcGIS · Kabul basin

Introduction

Water, which exists in nature in various forms, is the lifeblood of the world. But groundwater embraces a prominent function as a reliable source of fresh and potable water in most parts of the world. The water demand of 3 billion people worldwide is supplied from groundwater sources (Aggarwal et al., 2021). Despite the fact that groundwater is a part of the hydrological cycle, population growth and overextraction of groundwater have caused groundwater depletion in many Asian and North American nations (Vaiphei et al., 2020). Groundwater resources are increasingly vulnerable to overexploitation in arid

S. K. Singh · A. R. Noori (✉)
Department of Environmental Engineering, Delhi Technological University, Delhi, India
e-mail: a.noori@kpu.edu.af

A. R. Noori
Department of Water Supply and Environmental Engineering, Faculty of Water Resources and Environmental Engineering, Kabul Polytechnic University, Kabul, Afghanistan

Published online: 16 August 2022

and semi-arid climates, and the deteriorating quality of groundwater resources is also a significant issue (Goyal et al., 2021). As the capital of Afghanistan, with a total population of 4.45 million (NSIA, 2020), Kabul city is currently fulfilling all its water demand from groundwater (Shokory & Rabanizada, 2020). Therefore, it can be argued that groundwater is currently the sole body of freshwater willing to address Kabul's whole water demand. Recent studies, however, indicate a sharp decline in groundwater levels and deterioration of its quality in the city (Hamdard et al., 2020; Jawadi et al., 2020; Noori & Singh, 2021a, 2021b; Shokory & Rabanizada, 2020; Singh & Noori, 2022; Zahid et al., 2019; Zaryab et al., 2017, 2021; Zaryab et al., 2022a, b).

Destruction of groundwater and its quality, including all physicochemical indicators, heavy metals, toxic elements, and all significant ions, is a worldwide concern (Tiwari et al., 2017). Groundwater pollution may be of natural or anthropogenic origin. Natural factors include geochemical reactions in minerals (like F, NO_3, and As), the decomposition of organic compounds in soil, and the leaching of air particles by rainfall. Anthropogenic components, on the other hand, can include agricultural fertilizer waste, organic and chemical leachate from landfills, septic and sewage system penetration, and industrial chemical loss, among other things (Gupta et al., 2015; Paiman & Noori, 2019; Srinivas et al., 2015; Yadav et al., 2012). Groundwater quality degradation restricts its usage for household and commercial reasons. Once contaminated, the entire aquifer system will be impacted for thousands of years, attributed to the slow movement of groundwater (Brindha et al., 2011; Jerry, 1986).

As previously stated, groundwater has the main role for water supply in Kabul city. Therefore, it is essential to pay attention to its quality, regulate the quality attributes, and compare them to national and international standards to avoid quality degradation and preserve public health. It is critical to analyze groundwater quality in order to ascertain if it is suitable for human consumption. Water quality monitoring is essential for public health and the sustainability of water bodies. In this context, the Water Quality Index (WQI) methodology is an active instrument that provides policymakers and concerned individuals in the region with relevant and thorough information regarding water quality (Kanakiya et al., 2014; Raghav & Singh, 2021). ArcGIS has evolved into a robust tool for manipulating data, automating processes, and creating multivariate thematic maps to aid in geo-environmental decision-making (MacHiwal et al., 2011). Numerous investigations on groundwater quality and pollution owing to natural and artificial processes have been conducted (Arumugam & Elangovan, 2009; Babiker et al., 2007; Chatterjee et al., 2010; Duraisamy et al., 2019; Kawo & Karuppannan, 2018; Longe & Balogun, 2010; Magesh et al., 2013; Nas & Berktay, 2010; Sadat-Noori et al., 2014; Sarma & Singh, 2021; Shahid et al., 2014; Shomar et al., 2010; Singh et al., 2012; Tiwari & Singh, 2014; Vasanthavigar et al., 2010).

Since groundwater quality is affecting by alteration of land cover pattern, population growth, urbanization, and industrialization, water quality assessment as a monitoring tool is always necessary for frequent water quality monitoring. The current study aimed to develop a geospatial distribution model of water quality attributes in an intelligible method for the study region by integrating the water quality index (WQI) and geographic information system (GIS) with water quality statistics. The model might aid in determining the overall quality of groundwater and find out if there is any concern to a variety of groundwater consumers. Investigates and validates aquifers vulnerability, and demonstrate efficacy in conservation and rehabilitation initiatives. Furthermore, the study used physicochemical characteristics as well as significant anions and cations data to explore the hydrogeochemical mechanism of groundwater, determine the suitability of Kabul's groundwater for drinkable purposes, and find out the most unfavorable water sources. The findings of the study offer further insights regarding water quality assessment in the Kabul basin, and local authorities can utilize the outputs for sustainably development of groundwater management.

Material and methods

Study area

The research region is the Kabul Basin (JICA, 2011), which includes most parts of Kabul city, the center of Afghanistan. The research area is located in the country's central-east region, between 34°36′30" N and 34°24′40" N latitude and 69°01′25" E and 69°22′30" E longitude (Fig. 1). The area has a total size of 496 km^2. The climate condition of the study area is arid to

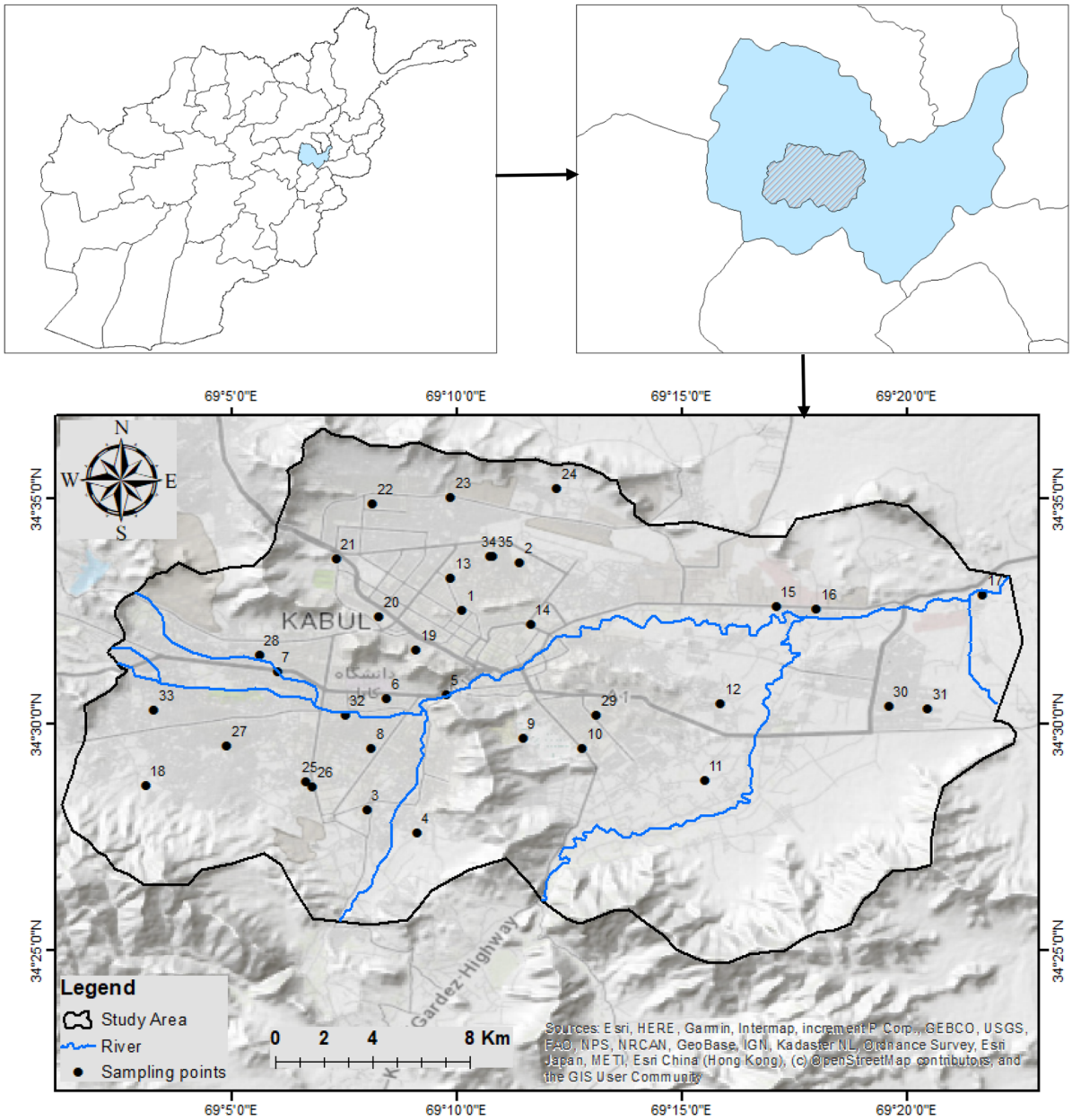


Fig. 1 Geographical location of the study area

semiarid. Rainfall in Kabul is seasonal, and usually, it is snowing and raining during winter (December, January, February) and early spring (March, April). The study area's annual average precipitation measured at around 330 mm per year (Source: Afghanistan meteorological department, data recorded 2008–2018). The highest and lowest average range of temperature in the basin is recorded as 32 °C in July and − 7 °C in January (Zaryab et al., 2017). The most common land use typology in the research region is built-up. It is the country's leading national commercial base, and many refugees returned to their homeland after 2001, which mostly settled in Kabul city. Also, a considerable number of people from other provinces came to Kabul to find a career. Low but quite steep mountain ranges enclose the basin. The elevation of the basin

ranges from 1763 to 2823 m (Fig. 2). Generally, the mountains in the southern and southwestern boundaries have a higher elevation than the others.

Drainage system

Three waterways have entered the city of Kabul. Paghman stream spilling out of the west to east. The Maidan River (Kabul River) arrive at the study area from the south and flows for 21 km before joining the Paghman River. The Logar river, an enormous tributary of the Kabul River, flows south-north and joins the Kabul River around 17 km downstream of the Paghman waterways mouth (Fig. 2). In recent years, the Kabul and Paghman rivers have some discharge only in late winter due to rainfall and the spring period due to snow melting and raining. The river beds have seemed to be the primary medium of groundwater recharge in Kabul. Groundwater recharge from precipitation through infiltrating accounts for a small percentage due to many reasons such as low precipitation, high evapotranspiration, covering the recharge zones by built-up areas, and clay layers. (Tünnermeier & Houben, 2005).

Geology

The research region is a part of the flat Kabul Basin. Metamorphic rocks make up the majority of the mountain range that surrounds and underpins the study region. The structure of the basin developed due to plate movements during the Late Paleocene time. At the early Tertiary (Late Paleocene) time, plate movements shaped the Kabul Basin's structure. The rocks, which are part of the Kabul block, are intersected by the Herat-Bamyan-Panjsher major fault in the west and north-west, the Sorobi fault in the east, and the Chaman fault system in the southeast. The common surficial geological forms

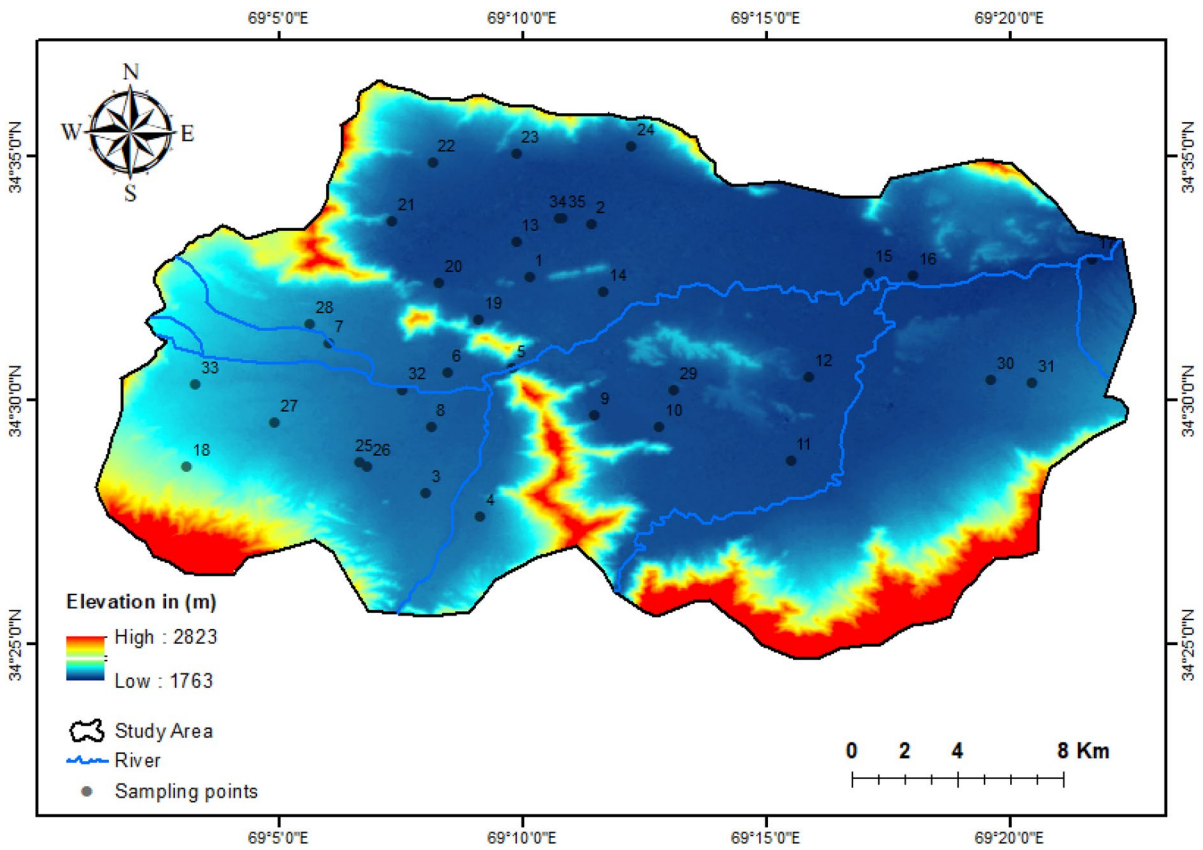


Fig. 2 Elevation and stream map of the basin

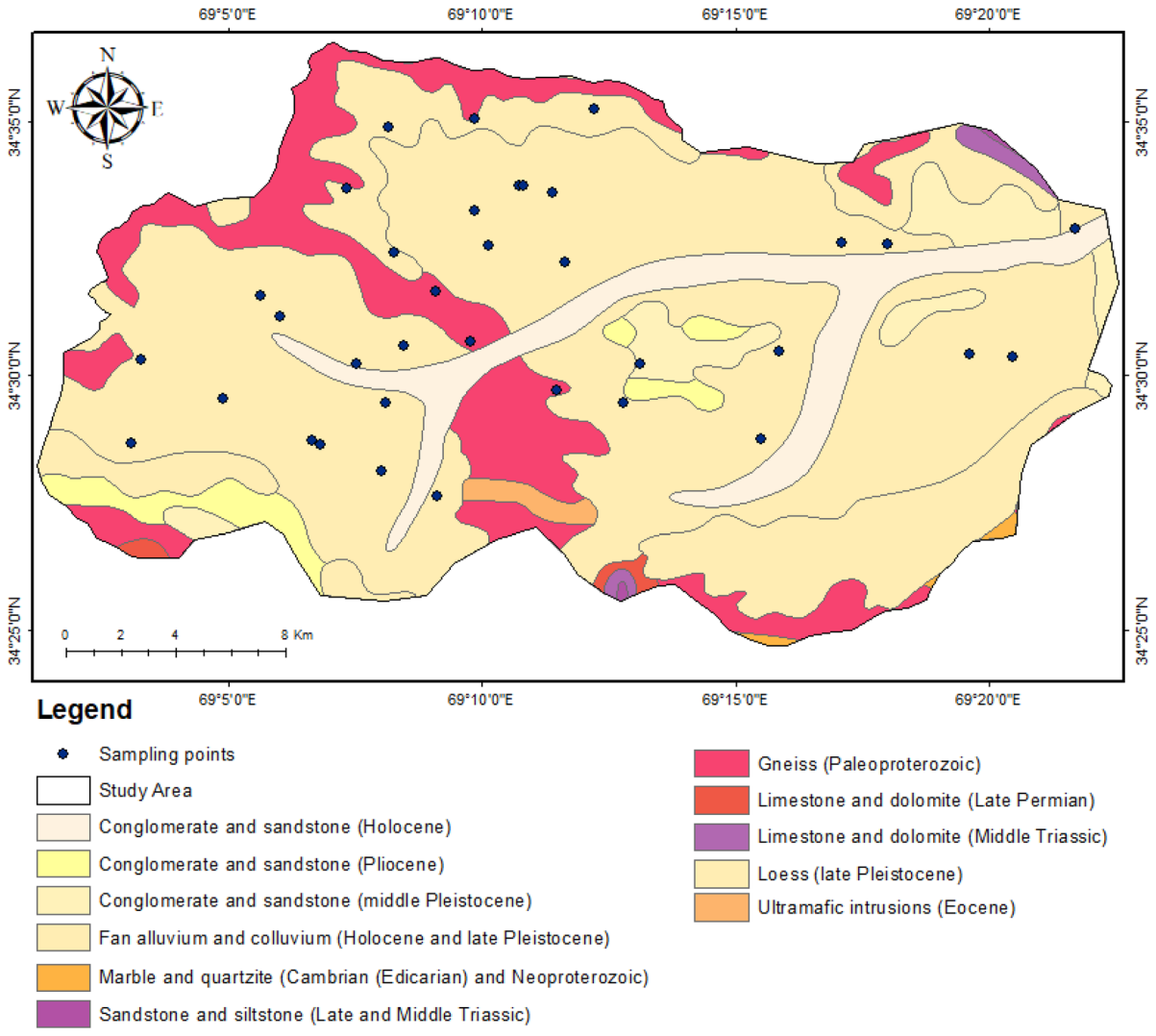


Fig. 3 Surficial geology of the basin

(Fig. 3) confirmed in the Kabul basin are conglomerate and sandstone, loess, metamorphic rocks, limestone, fan alluvium and colluvium, gneiss, limestone and dolomite, sandstone and siltstone, ultramafic intrusions, and river channel alluvium. The Kabul Basin is filled with an aggregation of terrestrial and lacustrine deposits, mostly uncemented and semi-consolidated lacustrine, fluvial, and eolian sedimentary rocks such as sand, gravel, and silt from the Quaternary and Neogene periods (Mack et al., 2009; Tünnermeier & Houben, 2005; Zaryab et al., 2017).

Hydrogeology

The hydrogeological condition of the Kabul basin is mainly formed of four interconnected Quaternary aquifers. The western region of the city, along the banks of the Paghman River and the upper stream of the Kabul River, has two aquifers. Along the Logar River path and the lower portion of the Kabul River course, two interconnected aquifers spanned the southeast, east, and northeast (lower Kabul subbasin). The Quaternary and Tertiary upper aquifers are mostly extracted for supplying water for the city.

The most commonly stated processes for groundwater recharge in Kabul city include recharge from riverbeds, recharging from snowmelts in the foothills, drainage channels, and direct recharge from precipitation surplus (Zaryab et al., 2017). Recharge from the water distribution network, wastewater filtration, septic system leaks, and toilet facilities are all possibilities. The high flow in the river is usually accompanied by the melting of snow and spring rains (April and May), which will most likely be the same period of high groundwater recharge.

Field and laboratory setup

Before sampling water from the hand pump and submersible wells, the water has been pumped out for 10 min. For the sample collection, 1-L prewashed HDPE bottles were utilized. The bottles have been washed with sample water 2–3 times. Thirty-five samples were acquired from different parts of Kabul city, with each municipality district receiving 2–3 samples to enable optimal spatial distribution and coverage of the whole research region. A portable digital multiparameter equipment (LABMAN Scientific instrument) was used to measure the temperature, pH, TDS, and EC on-site. Collected samples were moved to the water lab of Delhi Technological University's Department of Environmental Engineering and were promptly analyzed. Total hardness, chloride, and bicarbonate were evaluated utilizing the titrimetric technique (APHA, 2005). Sodium, calcium, magnesium, and potassium contents were measured by using a flame photometer. Fluoride was determined by using a digital portable multiparameter. The concentration of sulfate and nitrate was determined using UV–VIS spectrophotometers (UV 3092 and LT-290).

Water quality index

The Water Quality Index (WQI) is extensively regarded as the most comprehensive approach for assessing water quality (Rabeiy, 2018). Several studies have utilized the WQI to evaluate the quality and suitability of various types of water (Adimalla, 2021; Aenab et al., 2012; Aenab & Singh, 2013; Akter et al., 2016; Alum et al., 2021; Boyacioglu, 2007; Chakraborty et al., 2021; Chandrakant et al., 2021; Haritash et al., 2017; Karunanidhi et al., 2021; Khan et al., 2021; Liou et al., 2004; Mahmud et al., 2020; Oukil et al., 2021; Rabeiy, 2018; Radouane et al., 2021; Ram et al., 2021; Ramakrishnaiah et al., 2009; River et al., 2018; Shan et al., 2021; Tunc Dede et al., 2013; Uddin et al., 2021; Vaiphei et al., 2020; Vasanthavigar et al., 2010; Yisa & Jimoh, 2010; Zhang et al., 2021). In the current investigation, the WQI assesses groundwater quality according to 13 variables evaluated at each location. The weight value (w_i) for each metric varies from 1 to 5 depending on the degree of contamination's influence on human health (Table 1). At the first step of WQI computation, the relative weight of each parameter is calculated, as indicated in Eq. (1):

$$W_i = w_i / \sum_{i=1}^{n} w_i \quad (1)$$

where *Wi* stands for relative weight, *wi* for variable weight, and n for the number of variables. The quality rating scale (q_i) for each parameter will be calculated in the second step utilizing Eq. (2):

$$q_i = \frac{C_i}{S_i} \times 100 \quad (2)$$

where q_i is quality rating, and C_i is a measured concentration for each parameter of samples in mg/l. S_i is the standard limit for each parameter according to WHO in mg/l.

Table 1 Relative weights (w_i) for different water quality parameters

Chemical parameters		WHO Limits mg/lit	Weights *Wi*	Relative weights $Wi = wi/\sum wi$
1	TDS	1500	5	0.116
2	NO_3^-	45	5	0.116
3	F^-	1.5	5	0.116
4	pH	8.5	4	0.093
5	EC	1500	4	0.093
6	SO_4^{2-}	400	4	0.093
7	HCO_3^-	600	3	0.070
8	Cl^-	600	3	0.070
9	TH	500	3	0.070
10	Na^+	200	2	0.047
11	K^+	12	2	0.047
12	Ca^+	200	2	0.047
13	Mg^+	150	1	0.023
Total			43	1

By multiplying Eqs. (1) and (2), sub-indices are calculated:

$$SI_i = W_i \times q_i \quad (3)$$

Ultimately, WQI (Eq. 4) is calculated by summing all SI values for a particular sample, yielding a composite score based on ranges and water classes:

$$WQI = \sum SI_i \quad (4)$$

Geographic information system analysis

ArcGIS aids the analytical extrapolation of various experimental results to create theme maps and geospatial representations. It reduces the number of unknown water characteristics. It enables a statistical approach to graphically characterize groundwater quality in the research area. The most commonly used and recommended methods for generating spatial distribution maps are inverse distance weighting, kriging, and cokriging. As a result, for 35 samples of the research region in this study, spatial distribution models of all variables for groundwater quality were developed by using inverse distance weighted (IDW) interpolation method in ArcGIS-10.7 software (as shown in Fig. 4). The IDW interpolation technique estimates missing parameters concerning an interval, with the nearest point receiving greater weightage and decreasing as the distance increases. Furthermore, several researchers utilize this approach to create geographic distribution maps of various characteristics. Interpolation maps assist individuals and decision-makers in understanding groundwater quality by providing a generalized picture of hydrogeochemical processes and the drainage system of the studied region. This can aid subsequent water management, prevention, source of pollution, and groundwater modeling for future groundwater conservation.

Results and discussion

All physicochemical parameters of main anions and cations levels are provided in mg/l except pH and electrical conductivity (Table 2). The above mentioned table also indicates the descriptive statistics of the data. It also displays the number of samples that exceeded the allowed limit and the percentages of observations that exceeded the allowable limit in line with WHO criteria. The following are the details of the groundwater quality metrics that were investigated for drinking purposes.

Physicochemical parameters

Electrical Conductivity (EC)

This characteristic determines the quantity of current that an aqueous medium can convey. Based on the WHO standard, the highest permitted level for electrical conductivity is 1500 µS/cm. The detected EC ranges from 541 to 15,060 µS/cm. Based on (Todd, 1980); there are five primary EC groundwater quality classifications, as illustrated in Table 3. No sample had an excellent water type in the study region with an EC of less than 250 µS/cm. As shown in Fig. 4a, 17.14% of observations have a good water type (EC, 250–750 µS/cm), 62.86% have a permissible water type (EC, 750–2000 µS/cm), 5.71% have a doubtful water type (EC, 2000–3000 µS/cm), and 14.29% have an unsuitable water type ($EC > 3000$ µS/cm). As a result of this research, it is apparent that 20% of groundwater is unadvisable for residential usage.

pH

The pH values of water are the most significant and deciding component for its level of corrosivity, and this is due to water reacting with CO_2 in the underground to create carbonic acid. The interaction of CO_2, CO_3, and HCO_3, and their balance, determines the pH content of natural water. The pH measurements found in this study ranged between 6.6 and 7.7, which are within the ANSA (Afghanistan National Standard Authority) standard acceptable limits. Based on the findings, the pH value in the research location is not acidic.

The following equations illustrate the naturally changes of pH in groundwater before reaching the aquifer system:

1. Carbonic acid is formed when rainwater reacts with the atmosphere:

$$H_2O + CO_2 = H_2CO_3$$

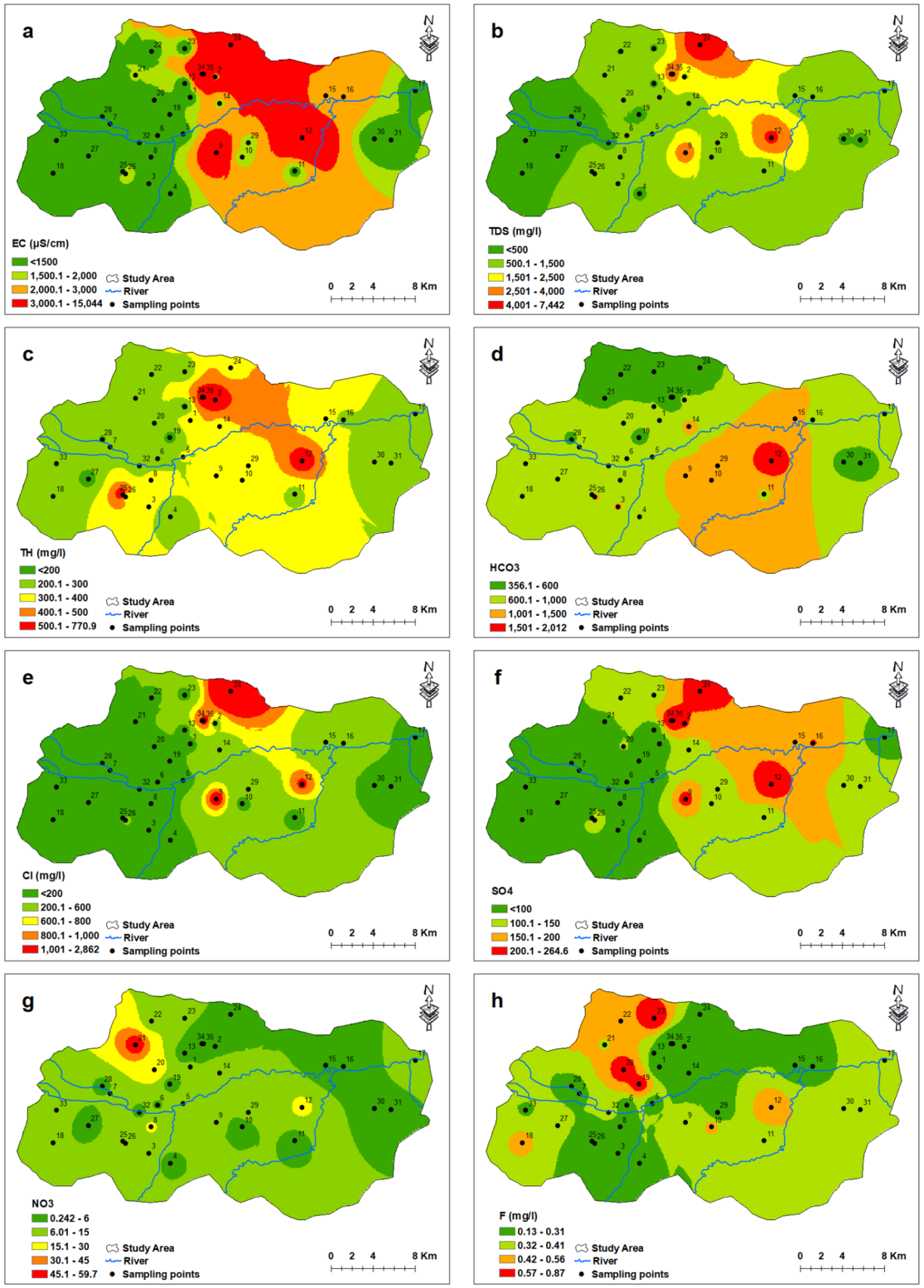


Fig. 4 Spatial distribution of observed parameters showing minimum to maximum ranges

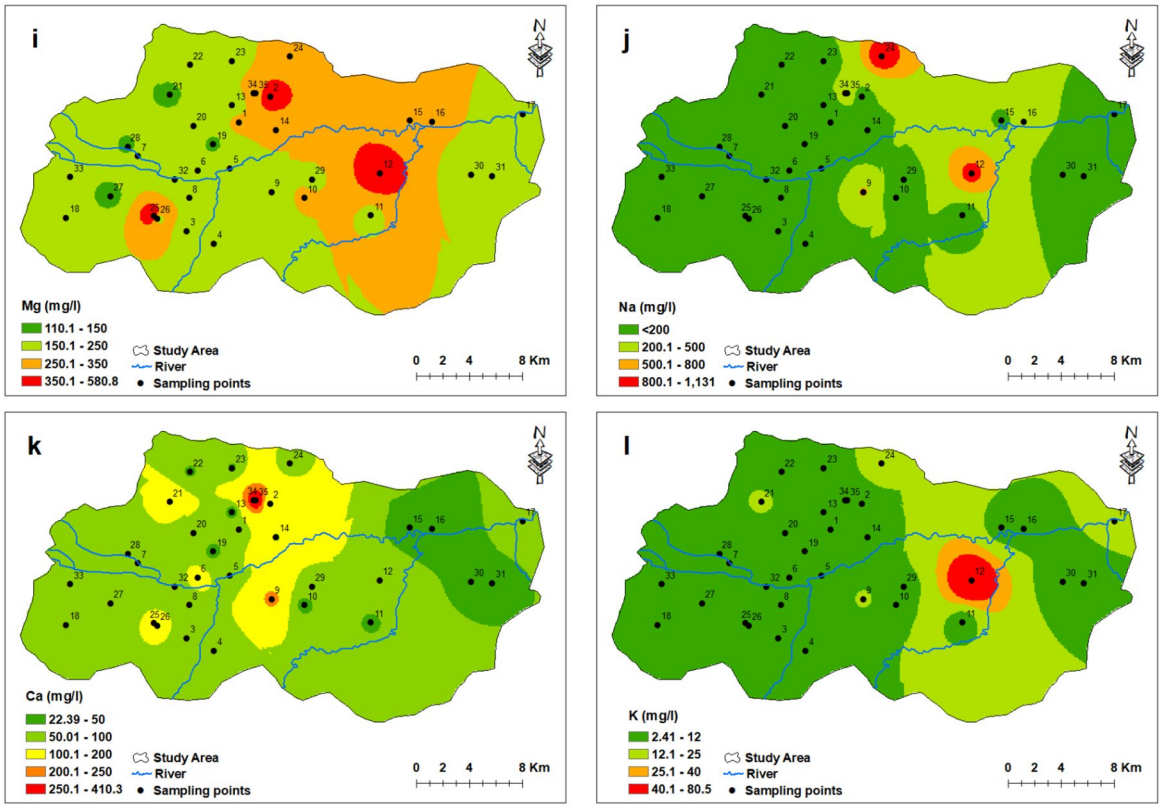


Fig. 4 (continued)

2. Carbonic acid dissociates into bicarbonate, releasing hydrogen ions, which make it acidic:

$$H_2CO_3 \rightarrow (HCO_3)^- + H^+$$

Table 2 Permissible limits of all parameters with reference to (WHO, 2011) standard showing minimum, maximum, average, median, and standard deviation values

Parameters	Min	Max	Mean	Median	SD	WHO limit	Samples exceeding permissible limit WHO	
							No. of samples	% of samples
EC	541.00	15,060.00	2280.00	1199.00	3050.52	750–1500	12.00	34.29
pH	6.60	7.70	7.19	7.20	0.19	6.5–8.5	Nill	Nill
TDS	270.00	7450.00	1135.20	598.00	1515.61	500–1500	5.00	14.29
TH	160.00	790.00	318.00	290.00	147.05	300–500	4.00	11.43
F	0.13	0.87	0.34	0.29	0.17	0.6–1.5	Nill	Nill
Cl	35.00	2865.00	315.86	100.00	557.36	250–600	5.00	14.29
NO_3	0.23	59.70	8.71	6.05	10.82	45.00	1.00	2.86
SO_4	22.87	264.76	114.07	102.79	74.61	250–400	Nill	Nill
HCO_3	353.80	2013.00	800.84	774.70	333.89	600.00	25.00	71.43
Na	23.60	1132.70	152.55	64.74	239.57	50–200	6.00	17.14
K	2.41	80.50	9.15	5.50	13.19	12.00	6.00	17.14
Ca	22.03	420.53	88.45	74.09	73.23	75–200	2.00	5.71
Mg	109.94	587.87	229.55	195.95	113.25	50–150	29.00	82.86

Table 3 Classification of groundwater according to electrical conductivity

EC	Classification	Samples exceeded permissible levels	% of sample that exceeds permissible levels
< 250	Excellent	Nill	Nill
250–750	Good	6	17.14
750–2000	Permissible	22	62.86
2000–3000	Doubtful	2	5.71
> 3000	Unsuitable	5	14.29
	Total	35	100.00

Total dissolved solids (TDS)

Total dissolved solids (TDS) refer to the inorganic salts and trace amounts of organic compounds found in aqueous solutions. Natural rock formation controls TDS primarily through weathering (transport), porosity, and permeability. Sewage disposal and fertilizer runoff are examples of anthropogenic sources. TDS readings with the highest and lowest range of 270–7450 mg/l were discovered, with an allowed limit of 2000 mg/l suggested by the ANSA standard. Figure 4b represents the geographical distribution layout of TDS in the research area. As indicated in Table 4, according to (Davis & DeWiest, 1966) classification, 40% of the samples are desirable for drinking with TDS less than 500 mg/l, 37.14% are permissible for drinking, having TDS 500–1000 mg/l, followed by 11.43% useful for irrigation, and 11.43% are unfit for drinking and irrigation. According to Freeze and Cherry (1979), 77.14% of observations are freshwater, 22.86% are brackish water, and no samples are saline or brine water, as indicated in Table 5.

Total hardness (TH)

Total hardness (TH) is an important factor in determining the properties of drinking water quality. It is reflecting the concentration of Ca and Mg ions in water. In the current study, TH concentrations vary from 160 to 790 mg/l. The total hardness maximum threshold has been limited to 500 mg/l by WHO guidelines. Water is divided into four categories based on its hardness (i.e., soft, moderate, hard, and very hard). No sample has been fallen in the soft and moderate water categories with TH less than 75 mg/l and (TH; 75–150 mg/l). As illustrated in Table 6 and the geographical distribution map (Fig. 4c), 57.14% of observations are in the hard group (TH; 150–300 mg/l), and 42.86% are in the very hard classification (TH > 300 mg/l).

Major ions chemistry

Anion concentrations were discovered in this investigation in the ascending order: $F < NO_3 < SO_4 < Cl < HCO_3$ with percentages of $0.03\% < 0.70\% < 9.20\% < 25.48$

Table 4 Groundwater classification based on total dissolved solid

TDS	Classification	Sample exceeded permissible levels	% of sample that exceeds permissible levels
< 500	Desirable for drinking	14	40.00
500–1000	Permissible for drinking	13	37.14
1000–3000	Useful for irrigation	4	11.43
> 3000	Unfit for drinking and irrigation	4	11.43
	Total	35	100.00

Table 5 Groundwater classification based on TDS content

TDS	Classification	Sample exceeded permissible levels	% of sample that exceeds permissible levels
< 1000	Fresh water type	27	77.14
1000–10,000	Brackish water type	8	22.86
10,000–100,000	Saline water type	Nill	Nill
> 100,000	Brine water type	Nill	Nill
	Total	35	100.00

% < 64.59% respectively. Bicarbonate (HCO_3^-) is the most common and stable ion found in water. Temperature, pH, dissolved CO_2, cations, and other salts all play a role in bicarbonate concentration in water. As shown in Fig. 4d, 71.4% of the samples are over the WHO standards, whereas 28.6% are below the permitted level. Bicarbonate levels in the Kabul basin ranging from 354 to 2013 mg/l. Chloride is the second most prevalent ingredient. Chloride in groundwater is due to weathering, sediment, soil leaching, and urbanization (Karanth, 1987). Cl^- ion concentrations in the examined water samples ranged between 35 and 2865 mg/l (Fig. 4e), with the exception of five samples, all being below the WHO permitted limit. Sulfate has a mean content of 114 mg/l and a variation of 22.87 to 265 mg/l, which is within the acceptable ranges (Fig. 4f). Sulfate is commonly found in groundwater as a result of the oxidation of sulfite minerals such as pyrite (FeS_2) (Yadav et al., 2012). The concentration of nitrate in the basin is within the acceptable limit (Fig. 4g) except for one sample that surpassed the WHO guidelines (greater than 45 mg/l). The region's average fluoride content is 0.34 mg/l, with minimum and maximum values ranging from 0.13 to 0.87 mg/l (Fig. 4h), all of which are within WHO's allowed thresholds. Fluoride source in groundwater is mainly of orogenic processes due to weathering of granitic rocks having F^--rich minerals as an accessory like amphiboles, apatite, fluorite, and mica (Reddy et al., 2010). Fluorosis of the teeth and skeleton is caused by high fluoride levels in water (beyond the permitted limit).

The quantity of cations is shown in ascending order: $K^+ < Ca^{+2} < Na^+ < Mg^{+2}$ contributing 1.91% < 18.44% < 31.80% < 47.85% respectively. Magnesium has the greatest proportion, with 229.54 mg/l on average (Fig. 4i). The highest amount of magnesium is 587.8 mg/l, and the lowest observation is 109.93 mg/l. Only 20% of the observed wells have a concentration within the allowable limit. The remaining 80% of the wells are above the WHO guidelines. Magnesium is found to have its source from all three types of rocks, such as igneous (basalt, dunites, pyroxenites), metamorphic (amphibolite, talc, tremolite-schists), and sedimentary (dolomite, gypsum) (Karanth, 1987). Sodium (Na^+) is the second component of the alkali metal group of studied groundwater. The sodium content in the Kabul basin groundwater ranges from 23.6 to 1132.7 mg/l (Fig. 4j). Comparing with WHO recommendations, 82.85% of the samples are within allowed ranges, whereas 17.14% of the detected samples exceed the permissible sodium limits. Na^+ is found in groundwater from silicate minerals. Rainwater, dissolution of vaporate minerals, sewage, and industrial outlet (Handa, 1975) are the other sources of sodium. With an average concentration of 88.45 mg/l,

Table 6 Classification of groundwater as per total hardness

TH as CaCO3	Classification	Sample exceeded permissible levels	% of sample that exceeds permissible levels
< 75	Soft	Nill	Nill
75–150	Moderately high	Nill	Nill
150–300	Hard	20	57.14
> 300	Very hard	15	42.86
	Total	35	100.00

calcium is the third most prevalent element in the sample. It constitutes 18.44% of the cation concentration. The highest calcium content is 420.52 mg/l, while the least amount is 22.02 mg/l (Fig. 4k). Only two of the 35 samples tested were found to be above the WHO's acceptable threshold. The main sources of calcium in groundwater are igneous and sedimentary rocks. Calcium and magnesium concentration relate to the hardness of water and are freely available on both surface, and subsurface water as carbonates and sources are from the rock like limestone, gypsum, and dolomite (Domenico & Schwartz, 1998). With a mean of 9.14 mg/l and a variation of 2.41 to 80.5 mg/l, potassium has the lowest concentration (Fig. 4l). The concentration of 17.14% of the samples in the basin exceeds the WHO recommendations, while the other samples are within limits. The distribution of different parameter concentration percentages in groundwater samples is depicted in Fig. 5. EC is first with 41.76%, TDS is second with 20.79%, HCO_3 is third with 14.67%, TH is fourth with 5.82%, and chloride is fifth with 5.79%.

Hydrochemical facies

The interaction between major anions and cations, as well as their behaviors, is addressed by the hydrochemical facies of groundwater. Hydrochemical facies aids in the identification and classification of various water types (Piper, 1944; Venugopal et al., 2009). As a result, the hydrogeochemical facies of groundwater

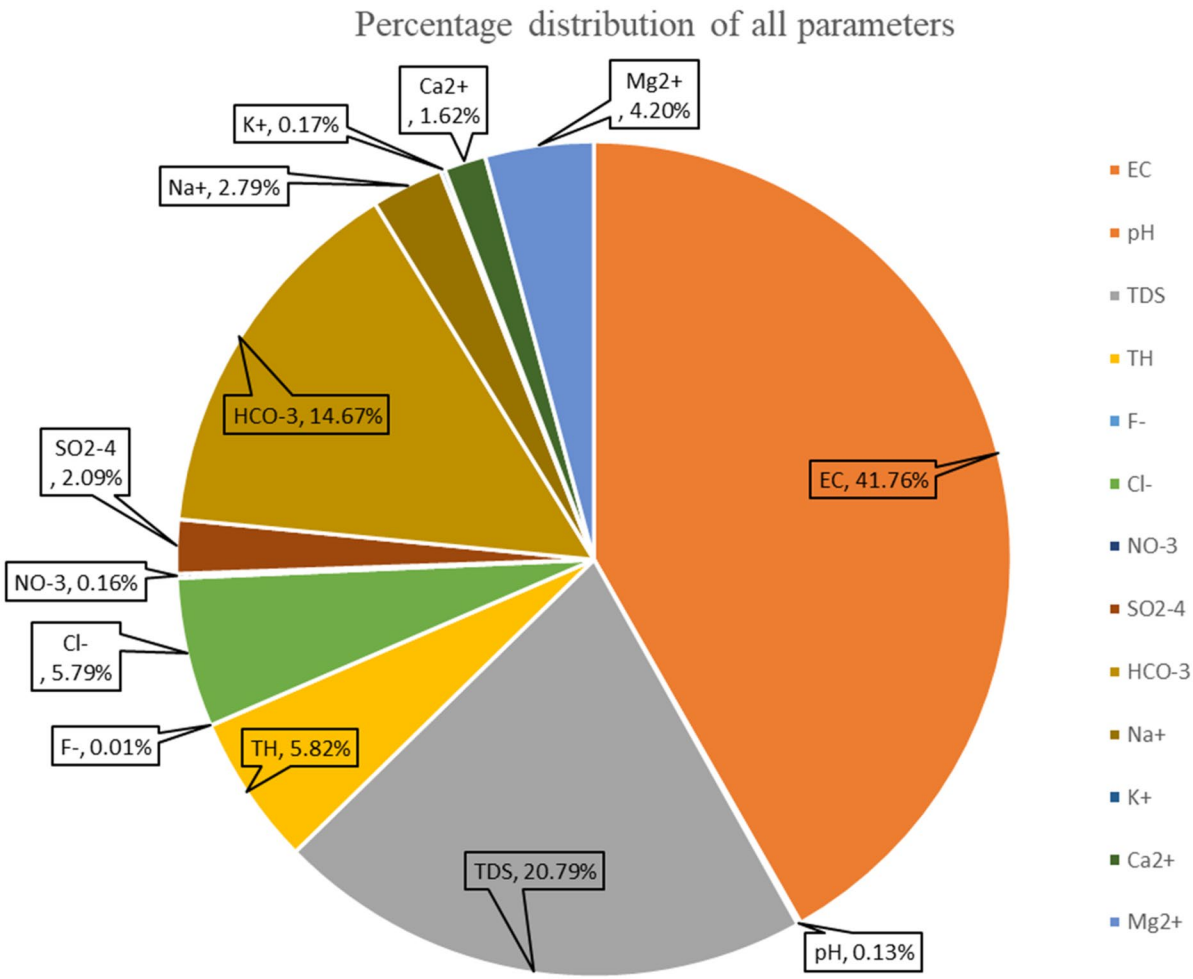


Fig. 5 Pie chart illustrating percentage distribution of all measured parameters

was estimated by plotting the concentrations of main anions (Cl, SO_4^{2}, and HCO_3) and cations (Ca^{2+}, Mg^{2+}, Na^{+}, and K^{+}) in meq/l in a Piper diagram (Fig. 6). There are six main water types in the geochemical evolution process. Six distinct forms of water categorization are depicted in Fig. 6. The following is the percentage distribution of samples of each kind of water: 45.71% mixed Ca-Na-HCO3 type, 37.14% Na-HCO3 type, 11.43% mixed Ca–Mg–Cl type, 2.86% each for Na-Cl type and Ca–Cl type, but 0 sample highlighted in Ca-HCO_3 type. Bicarbonate leads the proportion in the anions triangle with 82.85%, followed by chloride with 11.42%, and 5.71% of samples had no dominance. For cations, sodium is the most prominent ion with 80% dominance, no dominance with 17.14%, and 2.85% of samples with magnesium type.

Gibbs diagram

Gibbs plot illustrates the chemical specifications of water and the method of their interaction with various subsurface rock lithologies. The classification is of three types (Gibbs, 1970): type-I evaporation dominant (due to the rate of surface/subsurface evaporation), type-II rock dominance (due to chemical rock weathering by water), and type-III precipitation dominance (surface/subsurface precipitation). The Gibbs diagram was created by visualizing the main ions data from groundwater samples using the following two equations for anions and cations:

$$\text{Gibbs ratio I for anions}\left(\frac{\text{meq}}{\text{L}}\right) = \frac{\text{Cl}^-}{(\text{Cl}^- + \text{HCO}_3^-)} \quad (5)$$

$$\text{Gibbs ratio II for cations}\left(\frac{\text{meq}}{\text{L}}\right) = \frac{\text{Na}^+ + \text{K}^+}{(\text{Na}^+ + \text{K}^+ + \text{Ca}^{2+})} \quad (6)$$

As illustrated in Fig. 7, with the exception of some samples that are dominated by evaporation, the rock category dominates the bulk of the samples. The evaporation process is rare in both surface water and subsurface water as in common, but the hot and dry condition of the region's climate can cause that some of the samples fall in it (Selvakumar et al.,

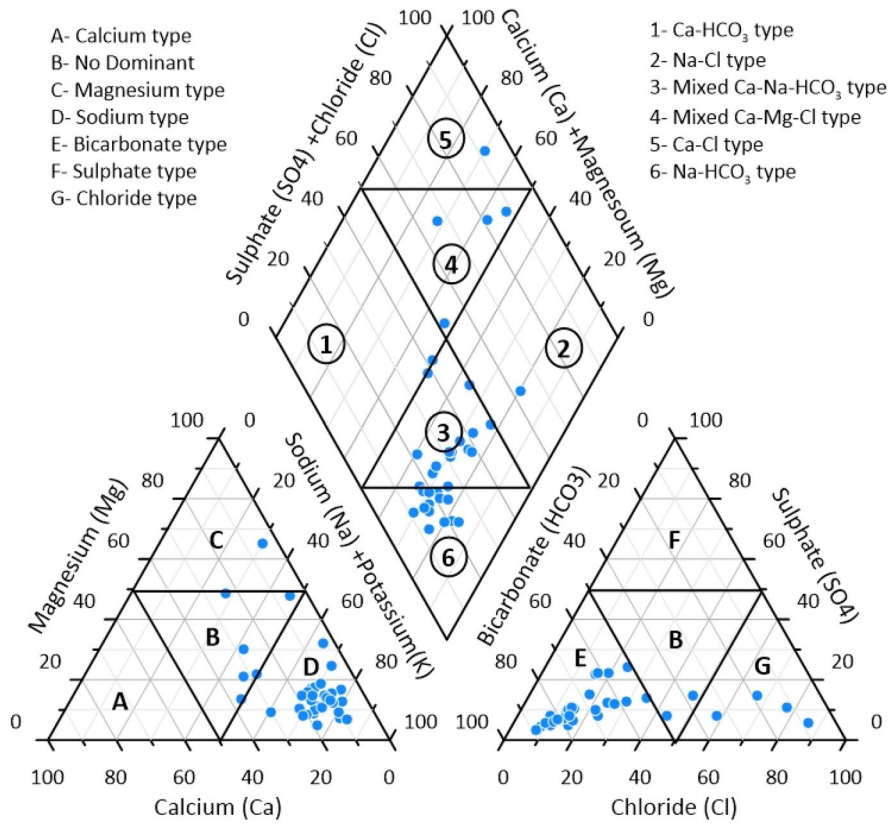


Fig. 6 Piper diagram illustrating various hydrochemical facies of groundwater

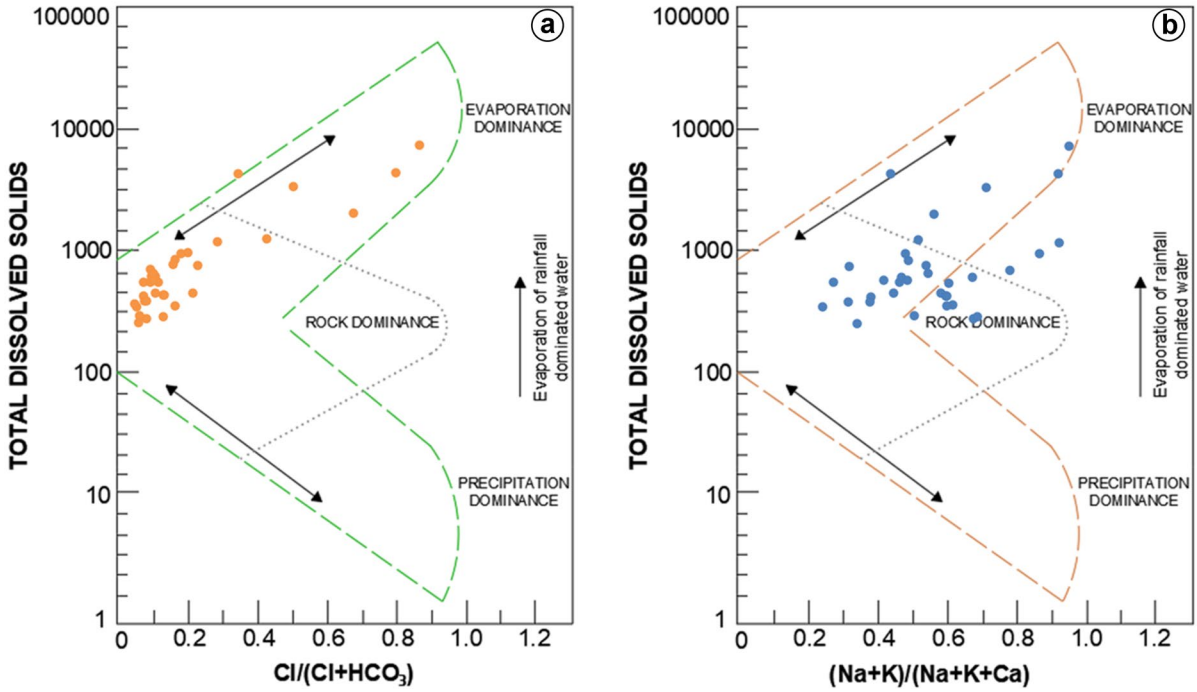


Fig. 7 Demonstration of principal influencing groundwater chemistry in Gibbs plot

2017). Nonetheless, maximal observations in rock domination show that silicate rocks are chemically degraded with groundwater at appropriate pressure and temperature.

Water quality index of groundwater

WQI is a concise, consistent, and dependable index for analyzing the quality of groundwater and determine its

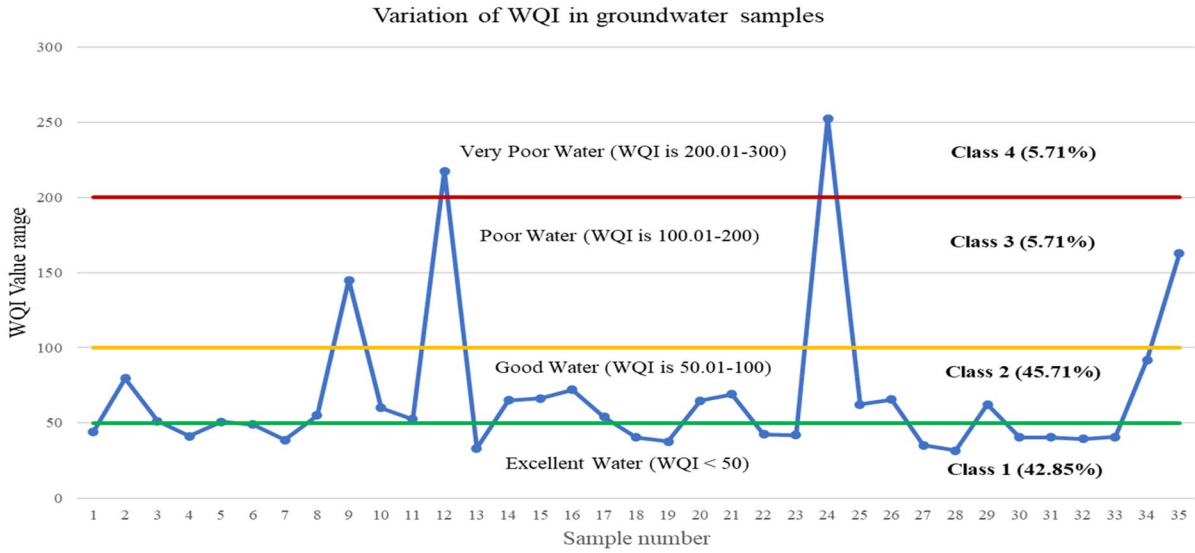


Fig. 8 Scatter plot showing water quality index (WQI) distribution in samples

Table 7 Water quality index-based groundwater classifications

WQI level	Water type	Samples	Percentage
< 50	Excellent water	15	42.86
50.00–100	Good water	16	45.71
100.01–200	Poor water	2	5.71
200.01–300	Very poor water	2	5.71
> 300	Unfit for drinking proposes	Nil	Nil
	Total	35	100.00

safety for drinking. Water has been categorized into five classes based on WQI values, including excellent water type (WQI less than 50), good water type (WQI 50–100), poor water type (WQI 100–200), very poor water type (WQI 200–300), and unsafe for drinking (WQI greater than 300). In the Kabul basin, most of the samples assessed were deemed to be good water. As shown in the scatter plot (Fig. 8), based on WQI results, 42.85% of observed samples are excellent water, and 45.71% of tested water is good water for drinking. The poor and very poor classes of water types each have 5.71%, respectively. As indicated in Table 7, no samples of the unfit for drinking class of water type have been identified. Figure 9 illustrates the WQI geospatial distribution map. The high value of WQI is seen in the central (wells 9 and 12) and northern parts (wells 24,34, and 35) of the basin.

Higher content of EC, chloride, sodium, and calcium, accompanied by sulfate and bicarbonate, indicates that in the investigated location, the rock–water interaction mechanism is the primary source of water quality decline. Agricultural activities may also have an impact on the sample wells in the center region of the basin. In the centre regions of the Kabul basin, there are several agricultural fields around the mentioned wells. The high level of WQI values in this area might be attributed to the widespread usage of fertilizers combined with a lack of appropriate waste disposal facilities.

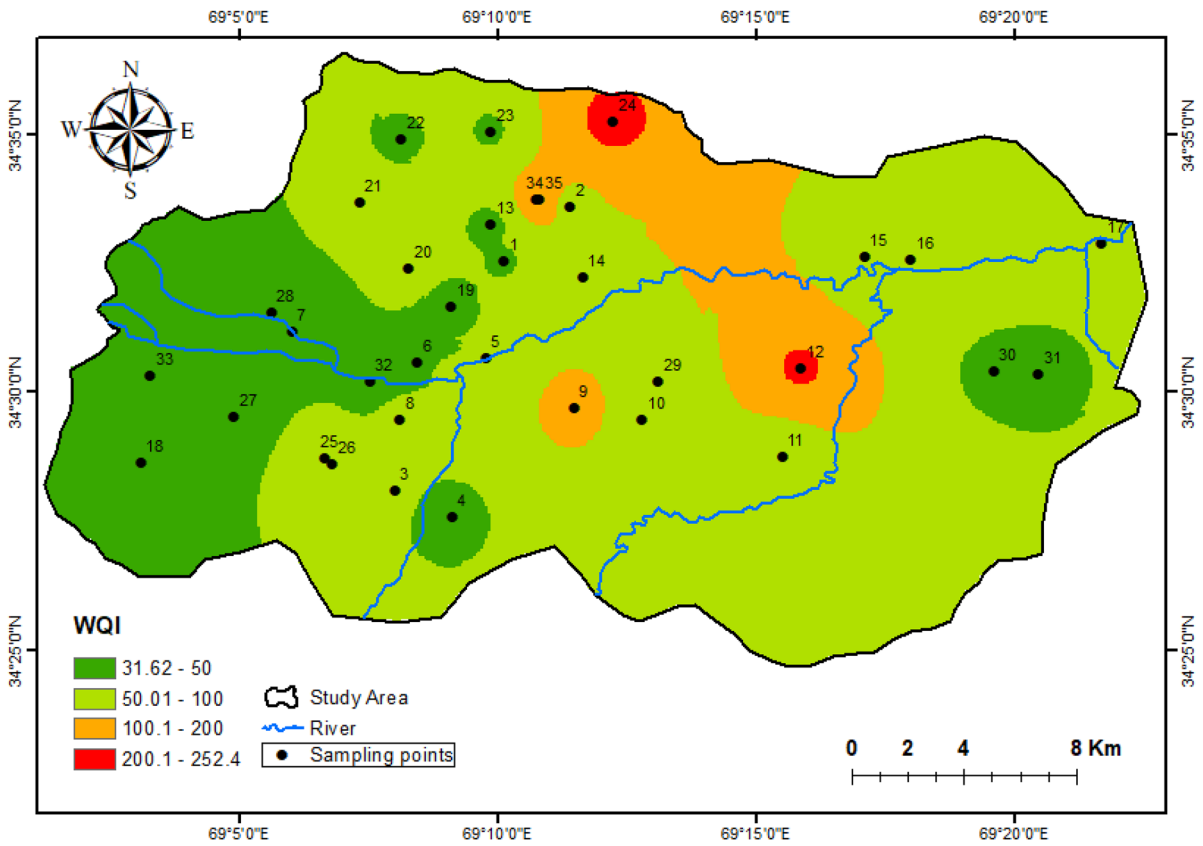


Fig. 9 Geospatial model of water quality index (WQI) in Kabul Basin

Conclusion

The research in Kabul comprises utilizing hydrogeochemical characterization, WQI, and GIS techniques to assess water quality in large parts of the city and most of its districts. According to WHO and ANSA guidelines, physicochemical characteristics such as EC, TDS, and TH are above acceptable levels in some samples. The quantity of anions in groundwater samples are in ascending order of $F < NO_3 < SO_4 < Cl < HCO_3$ with percentages of $0.03\% < 0.70\% < 9.20\% < 25.48\% < 64.59\%$ respectively. And the quantity of cations is shown in ascending order of $K < Ca < Na < Mg$ contributing $1.91\% < 18.44\% < 31.80\% < 47.85\%$ respectively. In observed samples of the basin, the major ions are HCO_3, Cl for anions, and Mg^{+2} and Na^{+} for cations. In certain samples, major anions Cl, SO_4, NO_3, and cations Na, Ca, and K in some samples and Mg in most samples have exceeded the permitted limit. As per the Piper diagram, the groundwater type is 45.71% mixed Ca-Na-HCO_3 and 37.14% Na-HCO_3, with sodium (80%) as the most prevalent cation. With 82.85%, bicarbonate has the highest percentage in the anions triangle. As per Gibbs plot conclusions, the bulk of the samples belonged to the rock dominance cluster. The outputs of examined wells in the river's vicinity indicate that the river does not influence their quality. This might be attributed to reduced groundwater levels and the natural treatment of recharge water during the groundwater recharging process. The poor water quality might be associated with lowering groundwater levels and reaching various geological formations, notably in the north of the research region. Furthermore, deterioration in groundwater quality might be ascribed to anthropogenic factors in the center sections of the basin due to higher groundwater levels. As per the water quality index (WQI), excellent to good water is found in 88.57% of the basin, whereas poor to very poor water is found in 11.43%, indicating pollution concentrations. As a result of the WQI readings, the majority of the groundwater sources are safe to drink. As groundwater is the most prevalent source of water supply, a continuous monitoring system will guaranty the sustainability of the drinking water quality and access of the residents to safe water and improved management with sustainable groundwater development.

Acknowledgements The authors are grateful to Delhi Technological University in New Delhi for enabling them to conduct this research.

Author contribution All authors contributed to the study conception and design. Sample collection, lab works, data analysis, and the first draft of the manuscript were performed by ARN. Revision and final comments were done by SKS. All authors read and approved the final manuscript.

Data availability Raw data were generated at Department of Environmental Engineering of Delhi Technological University. Derived data supporting the findings of this study are available from the corresponding author based on reasonable request.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Competing interests The authors declare no competing interests.

References

Adimalla, N. (2021). Application of the entropy weighted water quality index (EWQI) and the pollution index of groundwater (PIG) to assess groundwater quality for drinking purposes: A case study in a Rural Area of Telangana State, India. *Archives of Environmental Contamination and Toxicology, 80*(1), 31–40. https://doi.org/10.1007/s00244-020-00800-4

Aenab, A. M., & Singh, S. K. (2013). Evaluating water quality of Ganga River within Uttar Pradesh State by water quality index analysis using C ++ program. *Civil and Environmental Research*.

Aenab, A. M., Singh, S. K., & Al-Rubaye, A. A. M. (2012). Evaluation of Tigris River by water quality index analysis using C++ program. *Journal of Water Resource and Protection, 04*(07), 523–527. https://doi.org/10.4236/jwarp.2012.47061

Aggarwal, A., Soni, J., Sharma, K., Sapra, M., Chitrakshi, K., & O., & Haritash, A. K. (2021). Hydrogeochemical assessment of groundwater for drinking and agricultural use: A case study of rural areas of Alwar. *Rajasthan. Environmental Management, 67*(3), 513–521. https://doi.org/10.1007/s00267-020-01361-x

Akter, T., Jhohura, F. T., Akter, F., Chowdhury, T. R., Mistry, S. K., Dey, D., et al. (2016). Water quality index for measuring drinking water quality in rural Bangladesh: A crosssectional study. *Journal of Health, Population and Nutrition, 35*(1), 1–12. https://doi.org/10.1186/s41043-016-0041-5

Alum, O. L., Okoye, C. O. B., & Abugu, H. O. (2021). Quality assessment of groundwater in an agricultural belt in eastern Nigeria using a water quality index. *African Journal*

of Aquatic Science. https://doi.org/10.2989/16085914.2021.1882375

APHA. (2005). *Standard methods for the examination of water and wastewater, 20th*. American Public Health Association.

Arumugam, K., & Elangovan, K. (2009). Hydrochemical characteristics and groundwater quality assessment in Tirupur region, Coimbatore District, Tamil Nadu, India. *Environmental Geology, 58*(7), 1509–1520. https://doi.org/10.1007/s00254-008-1652-y

Babiker, I. S., Mohamed, M. A. A., & Hiyama, T. (2007). Assessing groundwater quality using GIS. *Water Resources Management, 21*(4), 699–715. https://doi.org/10.1007/s11269-006-9059-6

Boyacioglu, H. (2007). Development of a water quality index based on a European classification scheme. *Water SA, 33*(1), 101–106. https://doi.org/10.4314/wsa.v33i1.47882

Brindha, K., Rajesh, R., Murugan, R., & Elango, L. (2011). Fluoride contamination in groundwater in parts of Nalgonda District, Andhra Pradesh, India. *Environmental Monitoring and Assessment, 172*(1–4), 481–492. https://doi.org/10.1007/s10661-010-1348-0

Chakraborty, B., Roy, S., Bera, A., Adhikary, P. P., Bera, B., Sengupta, D., et al. (2021). *Geospatial assessment of groundwater quality for drinking through water quality index and human health risk index in an upland area of Chota Nagpur Plateau of West Bengal, India*. https://doi.org/10.1007/978-3-030-63422-3_19

Chandrakant, A., Vaishali, K., & Raut, S. (2021). An assessment of water quality index of Godavari river water in Nashik city, Maharashtra. *Applied Water Science, 11*(6), 1–11. https://doi.org/10.1007/s13201-021-01432-2

Chatterjee, R., Tarafder, G., & Paul, S. (2010). Groundwater quality assessment of Dhanbad district, Jharkhand, India. *Bulletin of Engineering Geology and the Environment, 69*(1), 137–141. https://doi.org/10.1007/s10064-009-0234-x

Davis, S. N., & DeWiest, R. J. M. (1966). *Hydrogeology*. John Wiley and Sons.

Domenico, P., & Schwartz, F. (1998). *Physical and chemical hydrogeology*. Wiley.

Duraisamy, S., Govindhaswamy, V., Duraisamy, K., Krishinaraj, S., Balasubramanian, A., & Thirumalaisamy, S. (2019). Hydrogeochemical characterization and evaluation of groundwater quality in Kangayam taluk, Tirupur district, Tamil Nadu, India, using GIS techniques. *Environmental Geochemistry and Health, 41*(2), 851–873. https://doi.org/10.1007/s10653-018-0183-z

Freeze, R. A., & Cherry, J. A. (1979). *Groundwater*. Prentice-Hall.

Gibbs, R. J. (1970). Mechanisms controlling world water chemistry. *Science, 170*(3962), 1088–1090.

Goyal, D., Haritash, A. K., & Singh, S. K. (2021). A comprehensive review of groundwater vulnerability assessment using index-based, modelling and coupling methods. *Journal of Environmental Management*. Elsevier Ltd. https://doi.org/10.1016/j.jenvman.2021.113161

Gupta, N., Kumar, K., & VinitKumar, Y. (2015). A review on current status of municipal solid waste management in India. *Journal of Environmental Sciences*, *37*, 206–217. https://doi.org/10.1016/j.jes.2015.01.034

Hamdard, M. H., Soliev, I., Xiong, L., & Kløve, B. (2020). Drinking water quality assessment and governance in Kabul: A case study from a district with high migration and underdeveloped infrastructure. *Central Asian Journal of Water Research*, *6*(1), 66–81. https://doi.org/10.29258/cajwr/2020-r1.v6-1/66-81.eng

Handa, B. K. (1975). Geochemistry and genesis of fluoride-containing ground waters in India. *Groundwater*.

Haritash, A. K., Mathur, K., Singh, P., & Singh, S. K. (2017). Hydrochemical characterization and suitability assessment of groundwater in Baga-Calangute stretch of Goa. *India. Environmental Earth Sciences, 76*(9), 1–10. https://doi.org/10.1007/s12665-017-6679-5

Jawadi, H. A., Sagin, J., & Snow, D. D. (2020). A detailed assessment of groundwater quality in future development. *Water*, 1–19. https://doi.org/10.3390/w12102890

Jerry, N. (1986). *Basic environmental technology: Water supply, waste management and pollution control*.

JICA. (2011). *The study on groundwater resources potential in Kabul basin in the Islamic republic of Afghanistan final report*. JICA.

Kanakiya, R. S., Singh, S. K., & Sharma, J. N. (2014). Determining the water quality index of an urban water body Dal Lake, Kashmir, India. *IOSR Journal of Environmental Science, Toxicology and Food Technology, 8*(12), 64–71. https://doi.org/10.9790/2402-081236471

Karanth, K. (1987). *Ground water assessment: Development and management*. Tata McGraw-Hill Education.

Karunanidhi, D., Aravinthasamy, P., Subramani, T., & Muthusankar, G. (2021). Revealing drinking water quality issues and possible health risks based on water quality index (WQI) method in the Shanmuganadhi River basin of South India. *Environmental Geochemistry and Health, 43*(2), 931–948. https://doi.org/10.1007/s10653-020-00613-3

Kawo, N. S., & Karuppannan, S. (2018). Groundwater quality assessment using water quality index and GIS technique in Modjo River Basin, central Ethiopia. *Journal of African Earth Sciences, 147*(June), 300–311. https://doi.org/10.1016/j.jafrearsci.2018.06.034

Khan, R., Saxena, A., Shukla, S., Sekar, S., & Goel, P. (2021). Effect of COVID-19 lockdown on the water quality index of River Gomti, India, with potential hazard of faecal-oral transmission. *Environmental Science and Pollution Research, 28*(25), 33021–33029. https://doi.org/10.1007/s11356-021-13096-1

Liou, S. M., Lo, S. L., & Wang, S. H. (2004). A generalized water quality index for Taiwan. *Environmental Monitoring and Assessment, 96*(1–3), 35–52. https://doi.org/10.1023/B:EMAS.0000031715.83752.a1

Longe, E. O., & Balogun, M. R. (2010). Groundwater quality assessment near a municipal landfill, Lagos, Nigeria. *Research Journal of Applied Sciences, Engineering and Technology, 2*(1), 39–44.

MacHiwal, D., Jha, M. K., & Mal, B. C. (2011). GIS-based assessment and characterization of groundwater quality in a hard-rock hilly terrain of Western India. *Environmental Monitoring and Assessment, 174*(1–4), 645–663. https://doi.org/10.1007/s10661-010-1485-5

Mack, T. J., Akbari, M. A., Ashoor, M. H., Chornack, M. P., Coplen, T. B., Emerson, D. G., et al. (2009). *Conceptual*

model of water resources in the Kabul Basin , Afghanistan. USGS.

Magesh, N. S., Krishnakumar, S., Chandrasekar, N., & Soundranayagam, J. P. (2013). Groundwater quality assessment using WQI and GIS techniques, Dindigul district, Tamil Nadu, India. *Arabian Journal of Geosciences, 6*(11), 4179–4189. https://doi.org/10.1007/s12517-012-0673-8

Mahmud, A., Sikder, S., & Chandra, J. (2020). Assessment of groundwater quality in Khulna city of Bangladesh in terms of water quality index for drinking purpose. *Applied Water Science, 10*(11), 1–14. https://doi.org/10.1007/s13201-020-01314-z

Nas, B., & Berktay, A. (2010). Groundwater quality mapping in urban groundwater using GIS. *Environmental Monitoring and Assessment, 160*(1–4), 215–227. https://doi.org/10.1007/s10661-008-0689-4

Noori, A. R., & Singh, S. K. (2021a). Status of groundwater resource potential and its quality at Kabul, Afghanistan : A review. *Environmental Earth Sciences, 80*(18), 1–13. https://doi.org/10.1007/s12665-021-09954-3

Noori, A. R., & Singh, S. K. (2021b). Spatial and temporal trend analysis of groundwater levels and regional groundwater drought assessment of Kabul, Afghanistan. *Environmental Earth Sciences*. https://doi.org/10.1007/s12665-021-10005-0

NSIA. (2020). Estimated population of Afghanistan 2020-21. Kabul, Afghanistan.

Oukil, A., Soltani, A. A., Boutaghane, H., Abdalla, O., Bermad, A., Hasbaia, M., & Boulassel, M. R. (2021). A surrogate water quality index to assess groundwater using a unified DEA-OWA framework. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-021-13758-0

Paiman, Z., & Noori, A. R. (2019). Evaluation of wastewater collection and disposal in Kabul City and its environmental impacts. *Modern Environmental Science and Engineering*, *5*(5), 451–458. https://doi.org/10.15341/mese(2333-2581)/05.05.2019/012

Piper, A. M. (1944). A graphic procedure in the geochemical interpretation of water analyses. *American Geophysical Union*, 914–928.

Rabeiy, R. E. S. (2018). Assessment and modeling of groundwater quality using WQI and GIS in Upper Egypt area. *Environmental Science and Pollution Research, 25*(31), 30808–30817. https://doi.org/10.1007/s11356-017-8617-1

Radouane, E. M., Chahlaoui, A., & Maliki, A. (2021). Assessment and modeling of groundwater quality by using water quality index (WQI) and GIS technique in meknes aquifer (Morocco). *Geology, Ecology, and Landscapes, 00*(00), 1–13. https://doi.org/10.1080/24749508.2021.1944797

Raghav, A., & Singh, S. K. (2021). Drinking water quality evaluation for groundwater of Delhi , India using GIS techniques. *Ecology, Environment and Conservation*.

Ram, A., Tiwari, S. K., Pandey, H. K., Chaurasia, A. K., Singh, S., & Singh, Y. V. (2021). Groundwater quality assessment using water quality index (WQI) under GIS framework. *Applied Water Science, 11*(2), 1–20. https://doi.org/10.1007/s13201-021-01376-7

Ramakrishnaiah, C. R., Sadashivaiah, C., & Ranganna, G. (2009). Assessment of water quality index for the groundwater in Tumkur taluk, Karnataka state, India. *E-Journal of Chemistry, 6*(2), 523–530. https://doi.org/10.1155/2009/757424

Reddy, A. G. S., Reddy, D. V., Rao, P. N., & Prasad, K. M. (2010). Hydrogeochemical characterization of fluoride rich groundwater of Wailpalli watershed, Nalgonda District, Andhra Pradesh, India. *Environmental Monitoring and Assessment, 171*(1–4), 561–577. https://doi.org/10.1007/s10661-009-1300-3

River, B., Fathi, E., Zamani, R., Rafat, A., & Bidaki, Z. (2018). Water quality evaluation using water quality index and multivariate. *Applied Water Science, 8*(7), 1–6. https://doi.org/10.1007/s13201-018-0859-7

Sadat-Noori, S. M., Ebrahimi, K., & Liaghat, A. M. (2014). Groundwater quality assessment using the water quality index and GIS in Saveh-Nobaran aquifer, Iran. *Environmental Earth Sciences, 71*(9), 3827–3843. https://doi.org/10.1007/s12665-013-2770-8

Sarma, R., & Singh, S. K. (2021). Simulating contaminant transport in unsaturated and saturated groundwater zones. *Water Environment Research, 93*(9), 1496–1509. https://doi.org/10.1002/wer.1555

Selvakumar, S., Chandrasekar, N., & Kumar, G. (2017). Hydrogeochemical characteristics and groundwater contamination in the rapid urban development areas of Coimbatore, India. *Water Resources and Industry, 17*(February), 26–33. https://doi.org/10.1016/j.wri.2017.02.002

Shahid, S. U., Iqbal, J., & Hasnain, G. (2014). Groundwater quality assessment and its correlation with gastroenteritis using GIS: A case study of Rawal Town, Rawalpindi, Pakistan. *Environmental Monitoring and Assessment, 186*(11), 7525–7537. https://doi.org/10.1007/s10661-014-3945-9

Shan, V., Singh, S. K., & Haritash, A. K. (2021). Evaluation of water quality and potential metal contamination in ecologically important Bhindawas bird sanctuary, India. *Applied Water Science, 11*(1), 1–9. https://doi.org/10.1007/s13201-020-01334-9

Shokory, J. A. N., & Rabanizada, E. (2020). Sustainable household water-saving and demand management options for Kabul City. *IOP Conference Series: Earth and Environmental Science, 511*(1), 13. https://doi.org/10.1088/1755-1315/511/1/012003

Shomar, B., Fakher, S. A., & Yahya, A. (2010). Assessment of groundwater quality in the Gaza Strip, Palestine using GIS mapping. *Journal of Water Resource and Protection, 02*(02), 93–104. https://doi.org/10.4236/jwarp.2010.22011

Singh, S. K., & Noori, A. R. (2022). Delineation of groundwater recharge potential zones for its sustainable development utilizing GIS approach in Kabul basin, Afghanistan. *Arabian Journal of Geosciences*, *15*(2). https://doi.org/10.1007/s12517-021-09410-3

Singh, V. K., Bikundia, D. S., Sarswat, A., & Mohan, D. (2012). Groundwater quality assessment in the village of Lutfullapur Nawada, Loni, District Ghaziabad, Uttar Pradesh, India. *Environmental Monitoring and Assessment, 184*(7), 4473–4488. https://doi.org/10.1007/s10661-011-2279-0

Srinivas, R., Bhakar, P., & Singh, A. P. (2015). Groundwater quality assessment in some selected area of Rajasthan, India using fuzzy multi-criteria decision making tool.

Aquatic Procedia, *4*(Icwrcoe), 1023–1030. https://doi.org/10.1016/j.aqpro.2015.02.129

Tiwari, A. K., & Singh, A. K. (2014). Hydrogeochemical investigation and groundwater quality assessment of Pratapgarh District, Uttar Pradesh. *Geological Society of India, 83*, 329–343.

Tiwari, A. K., Singh, A. K., Singh, A. K., & Singh, M. P. (2017). Hydrogeochemical analysis and evaluation of surface water quality of Pratapgarh district, Uttar Pradesh, India. *Applied Water Science, 7*(4), 1609–1623. https://doi.org/10.1007/s13201-015-0313-z

Todd, D. K. (1980). *Groundwater hydrology*. Wiley.

Tunc Dede, O., Telci, I. T., & Aral, M. M. (2013). The use of water quality index models for the evaluation of surface water quality: A case study for Kirmir Basin, Ankara, Turkey. *Water Quality, Exposure and Health, 5*(1), 41–56. https://doi.org/10.1007/s12403-013-0085-3

Tünnermeier, T., & Houben, G. (2005). *Hydrogeology of the Kabul Basin part I : Geology , aquifer characteristics , climate and hydrography*. BGR.

Uddin, M. G., Nash, S., & Olbert, A. I. (2021). A review of water quality index models and their use for assessing surface water quality. *Ecological Indicators, 122*, 107218. https://doi.org/10.1016/j.ecolind.2020.107218

Vaiphei, S. P., Kurakalva, R. M., & Sahadevan, D. K. (2020). Water quality index and GIS-based technique for assessment of groundwater quality in Wanaparthy watershed, Telangana, India. *Environmental Science and Pollution Research, 27*(36), 45041–45062. https://doi.org/10.1007/s11356-020-10345-7

Vasanthavigar, M., Srinivasamoorthy, K., Vijayaragavan, K., Rajiv Ganthi, R., Chidambaram, S., Anandhan, P., et al. (2010). Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. *Environmental Monitoring and Assessment, 171*(1–4), 595–609. https://doi.org/10.1007/s10661-009-1302-1

Venugopal, T., Giridharan, L., Jayaprakash, M., & Periakali, P. (2009). Environmental impact assessment and seasonal variation study of the groundwater in the vicinity of River Adyar, Chennai, India. *Environmental Monitoring and Assessment, 149*(1–4), 81–97. https://doi.org/10.1007/s10661-008-0185-x

WHO. (2011). *Guidelines for drinking-water quality* (Fourth Edi.).

Yadav, K. K., Gupta, N., Kumar, V., Arya, S., & Singh, D. (2012). Physico-chemical analysis of selected ground water samples of Agra city, India. *Recent Research in Science and Technology, 4*(11), 51–54.

Yisa, J., & Jimoh, T. (2010). Analytical studies on water quality index of river Landzu. *American Journal of Applied Sciences, 7*(4), 453–458. https://doi.org/10.3844/ajassp.2010.453.458

Zahid, S., Asmawi, M. Z., & Abdullah, M. F. (2019). Assessment of the groundwater quality and pollution in Kabul City, Afghanistan. *Advanced Journal of Technical and Vocational Education*, *3*(2), 20–26. https://doi.org/10.26666/rmp.ajtve.2019.2.4

Zaryab, A., Nassery, H. R., & Alijani, F. (2021). Identifying sources of groundwater salinity and major hydrogeochemical processes in the Lower Kabul Basin aquifer, Afghanistan. *Environmental Science: Processes and Impacts, 23*(10), 1589–1599. https://doi.org/10.1039/d1em00262g

Zaryab, A., Nassery, H. R., & Alijani, F. (2022a). The effects of urbanization on the groundwater system of the Kabul shallow aquifers, Afghanistan. *Hydrogeology Journal*, (0123456789). https://doi.org/10.1007/s10040-021-02445-6

Zaryab, A., Nassery, H. R., Knoeller, K., Alijani, F., & Minet, E. (2022b). Determining nitrate pollution sources in the Kabul Plain aquifer (Afghanistan) using stable isotopes and Bayesian stable isotope mixing model. *Science of the Total Environment, 823*, 153749. https://doi.org/10.1016/j.scitotenv.2022.153749

Zaryab, A., Noori, A. R., Wegerich, K., & Kløve, B. (2017). Assessment of water quality and quantity trends in Kabul aquifers with an outline for future drinking water supplies. *Central Asian Journal of Water Research*, *3*(2), 3–11

Zhang, J., Zhou, J., Zhou, Y., Zeng, Y., Ji, Y., Sun, Y., & Lei, M. (2021). Hydrogeochemical characteristics and groundwater quality assessment in the plain area of Yarkant River Basin in Xinjiang, P.R. China. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-021-12851-8

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Implementation of In-loop Filter based DSC-FLL Control Approach for Power Quality Improvement in Isolated Microgrid

[1]Sombir
Delhi Technological University
Delhi, India
sskundu46@gmail.com

[2]Madhusudan Singh
Delhi Technological University
Delhi, India
madhusudan@dce.ac.in

[3]Ashutosh K. Giri
Government Engineering College
Bharuch, Gujrat, India
ashu_ee2002@rediffmail.com

2022 IEEE 31st International Symposium on Industrial Electronics (ISIE) | 978-1-6654-8240-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIE51582.2022.9831712

Abstract:-—In this paper, a delayed signal cancellation based frequency locked loop (DSC-FLL) control approach is designed and implemented to provide improved disturbance rejection ability and rapid dynamic response with reduced steady state error. In an isolated microgrid (MG), two filtering stages are used in this control i) in-loop filter and ii) reduced order generalized integrator based FLL (ROGI-FLL). The DSC-FLL control with VSC provides voltage/frequency stabilization, active/reactive power compensation, and harmonics abatement of nonlinear loads. In this work, a solar photovoltaic (SPV) array is interfaced to the dc-link through a bidirectional converter (BDC), and a modified Perturb & Observe (P&O) technique is utilized for drawing peak power from a SPV. A Battery storage system (BSS) is utilized to maintain the power level between the renewable energy sources (RES's) and the load under intermittent condition of RES's and unbalanced loading condition. The BDC control maintains the charging and discharging of the BSS and regulates the dc-link voltage. Moreover, the BDC also removes the second harmonic from the battery charging current. The dynamic response of isolated MG is analyzed under different operating conditions.

Keywords- DSC-FLL, BDC, MPPT, isolated MG, THD

I. INTRODUCTION

The developing energy interest with exhausting customary energy sources and rising ecological worries have constrained the world to search for another energy sources to address energy security concerns while sticking to a low carbon development system. The commitment of environmentally friendly power has expanded in the general energy blend of the world, which can decrease CO_2 outflows universally. A wind driven SEIG plays a major role in the field of power generation for standalone applications because of their rugged and brushless construction, reduced size, less maintenance cost, absence of synchronizing equipment and better transient response with inherent short circuit capability [1, 2]. To stabilize the terminal voltage and frequency of the wind driven SEIG in a prescribed range, an adaptive control algorithm with VSC is implemented [3]. A VSC with suitable controls provides power support, harmonics abatement, load leveling and power quality improvement [4]. A BDC is utilized to draw peak power from a SPV coupled at dc-link of the VSC [5]. An incremental conductance (IC) [6. 7] and a Perturb & Observe [8] method based peak power point tracking (PPPT) control approach are implemented to harness maximal power from a SPV. A BSS absorbs the surplus power under dynamic conditions and maintains the power balance between the power generation source and load [9]. Different phase locked loop (PLL) controls with VSC are described to minimizing power quality problems [10, 11] and a differentiation based PLL is utilized to improve the dynamic response with reduced error of the isolated MG [12]. The PLL, which has nearly reached a mature level of improvement in electricity and power applications in 1-phase and 3-phase systems, the FLL is yet to explore fully. The implementation of FLLs is stationary reference frame which makes their modeling, tuning, and overall better [13]. The FLLs gives rapid response with reduced steady state error, dc offset removal under nonlinear load and stabilizes the parameter of islanded MG [14, 15]. A linear time invariant (LTI) based time delay (TD)-PLL is implemented to obtain higher accuracy and rapid dynamic response [16]. PLL and FLL based control algorithms are described to estimate the phase angle, fundamental component and removal of dc-component in 3-phase grid connected system [17]. A VCO-less PLL control is utilized to reduce the computational burden of trigonometric operations and to provide a fast dynamic response under variable conditions [18]. A modified AANF control approach is utilized with dc-offset dismissal ability from the given output and rapid frequency evaluation, which permit the modified AANF control approach to quickly and precisely compute the fundamental components of the unbalanced and nonlinear load current [19].

The following are the main contributions of this paper:

1) The DSC-FLL control approach is utilized to compute the fundamental component of load current and phase angle under dynamic conditions of isolated MG with rapid convergence response along with reduced steady-state error under distorted load current.
2) Moreover, the DSC-FLL control for VSC provides multifunction's such as stabilization of frequency and voltage at common coupling point (CCP), power compensation, harmonics abatement and improved over all power quality.
3) The BDC maintains the charging and discharging current of BSS and regulate the dc-link voltage, where the BSS is also utilized to consume excess power and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:25:23 UTC from IEEE Xplore. Restrictions apply.

also supports the load during intermittency of wind energy and maintains the power equilibrium at CCP.

II. System Configuration

The proposed isolated MG is depicted in Fig. 1. The proposed MG consists of wind driven SEIG, a SPV with modified P&O MPPT control, battery energy storage, interfacing inductors, ripple filter, VSC and nonlinear load. A BDC is interfaced at dc-link to control the battery charging and maintain the dc-link voltage. The interfacing inductors are used to connect the VSC at CCP in an isolated MG [14]. The interfacing inductors are utilized to minimize the ripple currents in the VSC. To reduce the high frequency switching noise from the CCP voltages ripple filters are being used [15].

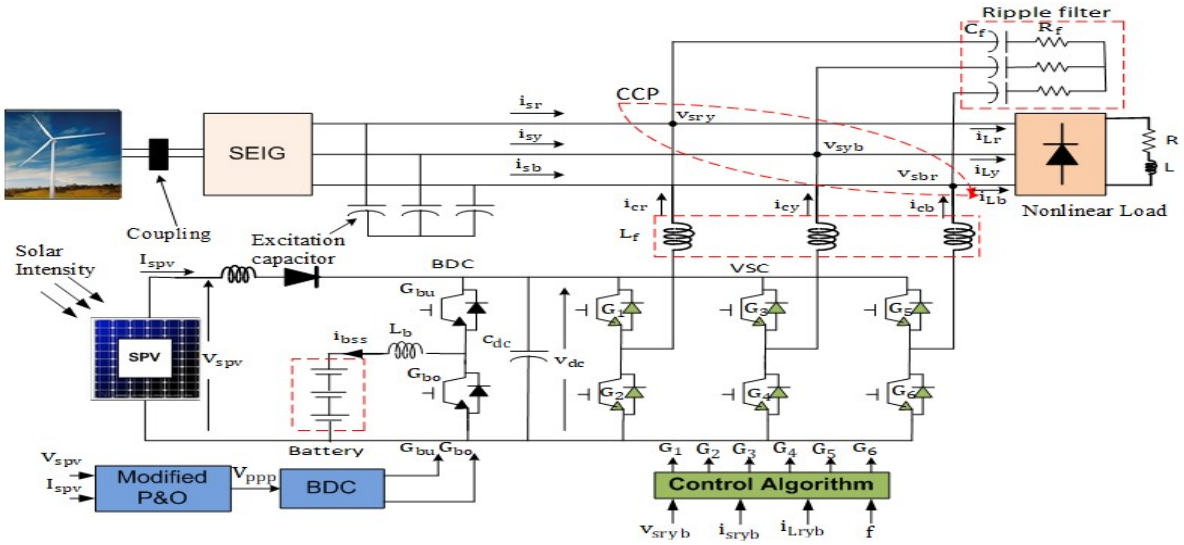


Fig. 1.Topology of isolated MG

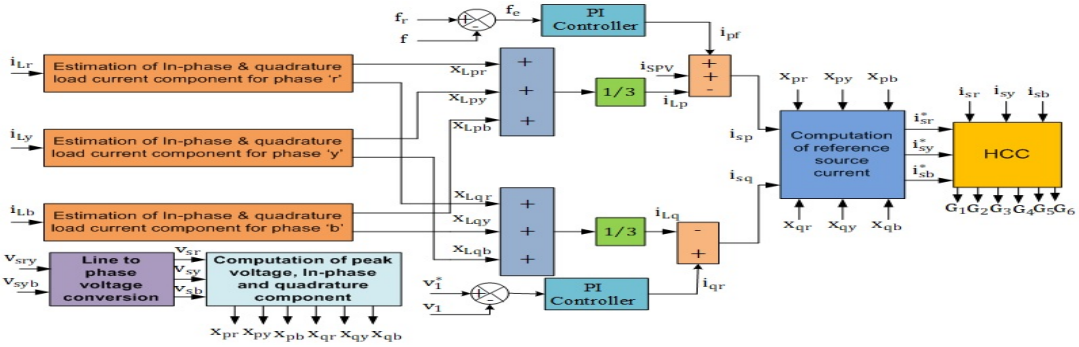


Fig. 2. Control scheme

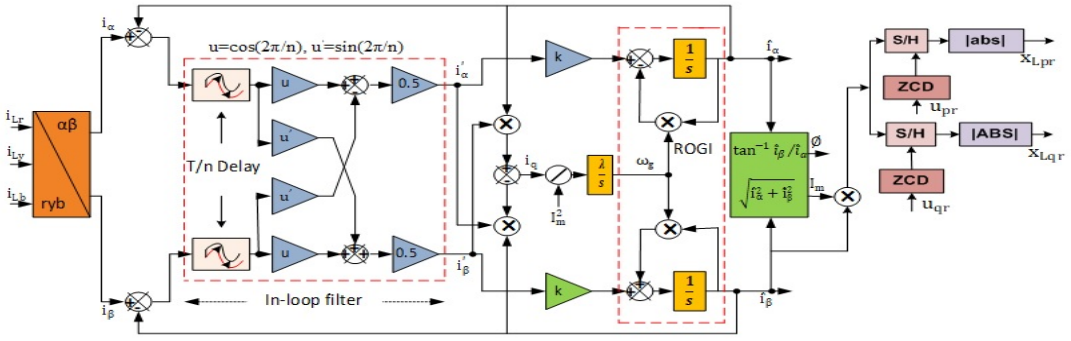


Fig.3. Control approach

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:25:23 UTC from IEEE Xplore. Restrictions apply.

III. DSC-FLL CONTROL ALGORITHM

The proposed system is used in standalone applications for supplying power in 1-phase and 3-phase unbalanced nonlinear loads. The DSC-FLL based controller for VSC builds the source currents sinusoidal and balanced while feeding the dynamic nonlinear loads. Moreover, the proposed control with the help of BSS is also utilized to compensate the dynamic power of the SPV and wind. It enables the utilization of renewable energy sources maximally for supplying the undisturbed power to the 1-phase or 3-phase loads for isolated locations. The control of isolated microgrid is categorized as (i) BDC with SPV peak power point tracking (PPPT) control, (ii) VSC control algorithm [18].

A. *MPPT and BDC Control*

In this isolated MG, a modified P&O control is utilized for implementation of MPPT. The P&O control output is employed for generating the duty cycle for the converter operating the SPV at its PPP. The duty cycle (α) for the converter is estimated as

$$\alpha = 1 - \frac{v_{ppp}}{v_{dc}} \tag{1}$$

Where, v_{ppp} and v_{dc} is the PPPT control output and battery voltage or dc-link voltage.

The fluctuations in dc-link voltage are due to variation in solar intensity level and unbalanced loading. A BDC control is utilized to regulate the dc-link voltage properly, and balance the CCP voltage and CCP frequency under load disturbances. The BDC is depending on the two cascaded PI controllers. The reference battery current is estimated by the inner PI controller. The equation of outer PI controller for voltage regulation is expressed as

$$i_{bss}(q) = i_{bss}(q-1) + k_p\{v_{dce}(q) - v_{dce}(q-1)\} + k_i v_{dce}(q) \tag{2}$$

$$\text{And, } v_{dce}(q) = v_{ppp}(q-1) - v_{dc}(q) \tag{3}$$

Where, v_{ppp} is the reference dc-link voltage evaluated through a modified P&O algorithm and v_{dc} is the sensed dc-link voltage. The battery current is controlled by the inner PI controller. The outer PI controller output is taken as reference battery current of the inner PI controller. The equation for the inner PI controller is described as

$$\alpha(q) = \alpha(q-1) + k_p\{i_{bsse}(q) - i_{bsse}(q-1)\} + k_i i_{bsse}(q) \tag{4}$$

$$i_{bsse}(q) = i^*_{bss}(q) - i_{bss}(q) \tag{5}$$

B. *VSC Control*

A DSC-FLL control is utilized to the control the VSC. The in-loop filter based DSC-FLL provides improved disturbance rejection ability. The concept of in-loop filters in FLLs is described as the computed frequency (ω_g) is assumed constant. The ROGI implementation in space vector notation is shown in fig. 3 expressed as

$$\hat{i}_\alpha(s) + j\hat{i}_\beta(s) = \frac{k}{s - j\omega_g + k}\left(i_\alpha(s) + j i_\beta(s)\right) \tag{6}$$

The proposed control passes the FFPS component and blocks the other frequency components. When extracting the FFPS component and eliminate the disturbances such as voltage imbalance and harmonics, the equation is expressed as

$$DSC_n = \frac{1 + e^{\frac{j2\pi}{n}} e^{-\frac{Ts}{n}}}{2} \tag{7}$$

The components (T and n) are the fundamental period and the operator delay factor respectively. Mainly, the capability of filtering depends on its delay factor (n). Also described in the literature, a single operator may not be capable to attenuate all the disturbances, so the operators are utilized in cascade form with different delay factor to remove such disturbances as shown in fig. 3. The equation for open loop transfer function in DSC-FLL can be expressed as

$$G^\theta = \frac{\theta_1(s)}{\theta_1(s) - \hat{\theta}_1(s)} = \frac{1 + e^{\frac{-Ts}{2}}}{2} \frac{1 + e^{\frac{-Ts}{12}}}{2} \frac{ks + \lambda}{s^2} \tag{8}$$

The constant parameters k and λ are evaluated as

$$k = \frac{1}{gT_d} \text{ and } \lambda = \frac{1}{g^3 T_d^2} \tag{9}$$

Where k=160, λ=12791 and g is the phase margin evaluation factor. The DSC-FLL based control accurately estimates the fundamental component and phase angle under dynamic conditions.

$$I_m = \sqrt{\hat{i}_\alpha^2 + \hat{i}_\beta^2} \tag{10}$$

$$\emptyset = \tan^{-1}(\hat{i}_\beta / \hat{i}_\alpha) \tag{11}$$

The magnitude of the fundamental active and quadrature components of the load currents is computed using S&H (Sample and Hold), ZCD (Zero Crossing Detector) and quadrature unit templates as depicted in Fig. 3. Similarly, the active and quadrature components of the leftover two load phase currents are computed using the same pattern. After the calculation of fundamental component current of load, two proportional integral (PI) controllers are utilized for regulation of voltage/frequency, and results of the PI controllers are deducted from the current of fundamental components of the load [17]. For the estimation of PI controller's current, the control equations are expressed as

$$i_{fp}(q) = i_{fp}(q-1) + k_{fp}\{f_e(q) - f_e(q-1)\} + k_{fi} f_e(q) \tag{12}$$

$$i_{qr}(q) = i_{qr}(q-1) + k_{vp}\{v_e(q) - v_e(q-1)\} + k_{vi} v_e(q) \tag{13}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:25:23 UTC from IEEE Xplore. Restrictions apply.

Where the frequency error and voltage error are estimated as

$$f_e(q) = f_r(q) - f(q) \text{ and } v_e(q) = v_1^*(q) - v_1(q) \quad (14)$$

The peak value of source voltage (v_1) from the CCP is computed as

$$v_1 = \sqrt{\left(\frac{2}{3}\right)} \times \sqrt{\left(v_{sr}^2 + v_{sy}^2 + v_{sb}^2\right)} \quad (15)$$

$v_{sr} = \frac{1}{3}(2v_{sry} + v_{syb})$, $v_{sy} = \frac{1}{3}(-v_{sry} + v_{syb})$, and $v_{sb} = \frac{1}{3}(-v_{sry} - 2v_{syb})$ (16)

Where, v_{sry} and v_{syb} are the line voltages of the source at CCP and v_{sr}, v_{sy} and v_{sb} are the computed phase voltages. The both active (i_{sp}) and reactive (i_{sq}) reference source currents are evaluated as

$$i_{sp} = i_{fp} + i_{SPV} - i_{Lp} \quad (17)$$

$$i_{sq} = i_{Lq} - i_{qr} \quad (18)$$

The active and reactive component of load current are evaluated with averaging three components

$$i_{Lp} = \frac{x_{Lpr} + x_{Lpy} + x_{Lpb}}{3} \text{ and } i_{Lq} = \frac{x_{Lqr} + x_{Lqy} + x_{Lqb}}{3} \quad (19)$$

The sinusoidal reference currents from load current is estimated with in-loop filter based DSC-FLL as follow

$$i_{sra}^* = i_{sp} \times x_{pr}, i_{syb}^* = i_{sp} \times x_{py}, i_{sbc}^* = i_{sp} \times x_{pb} \quad (20)$$

$$i_{qra}^* = i_{sq} \times x_{qr}, i_{qyb}^* = i_{sq} \times x_{qy}, i_{qbc}^* = i_{sq} \times x_{qb} \quad (21)$$

Where, x_{pr}, x_{py}, x_{pb}, x_{qr}, x_{qy} and x_{qb} are the estimated active and reactive unit templates. The evaluation of unit templates is described as

$$x_{pr} = {}^{v_{sr}}/_{v_1}, x_{py} = {}^{v_{sy}}/_{v_1}, x_{pb} = {}^{v_{sb}}/_{v_1} \quad (22)$$

$$x_{qr} = -\frac{x_{pr}}{\sqrt{3}} + \frac{x_{pb}}{\sqrt{3}}, x_{qy} = \frac{\sqrt{3}x_{pr}}{2} + \frac{(x_{py} - x_{pb})}{2\sqrt{3}}$$

$$\text{and } x_{qb} = -\frac{\sqrt{3}x_{pr}}{2} + \frac{(x_{py} - x_{pb})}{2\sqrt{3}} \quad (23)$$

The reference source currents is summation of extracted active and reactive current is expressed as

$$i_{sr}^* = i_{sra}^* + i_{qra}^*$$

$$i_{sy}^* = i_{syb}^* + i_{qyb}^* \quad (24)$$

$$i_{sb}^* = i_{sbc}^* + i_{qbc}^*$$

The difference of estimated reference current and source current is given to hysteresis controller for the switching pulses generation [8].

IV. Results and Discussion

An in-loop filter based DSC-FLL control approach is designed to reduce the power quality problems of an isolated MG under wind/solar intermittency and nonlinear load perturbation. The proposed control successfully extracts the unit vectors and generates the appropriate gating pulses for VSC.

A. Dynamic response of MG with changes in wind velocity

Fig. 4 shows the dynamic response of isolated MG under variations in wind velocity and constant solar intensity level. The wind velocity is changed from 14 m/s to 11 m/s at t=3.2s, the solar intensity and temp. (T) are fixed at 1000 w/m^2 and 35°C. During the change in wind speed the SEIG based source power is reduced from 5.65 kW to 5.49 kW. The DSC-FLL with VSC stabilizes the source voltage (v_{sryb}) and maintains source current (i_{sryb}). The compensator current is slightly increasing at t=3.2s and provides compensation at CCP through VSC.

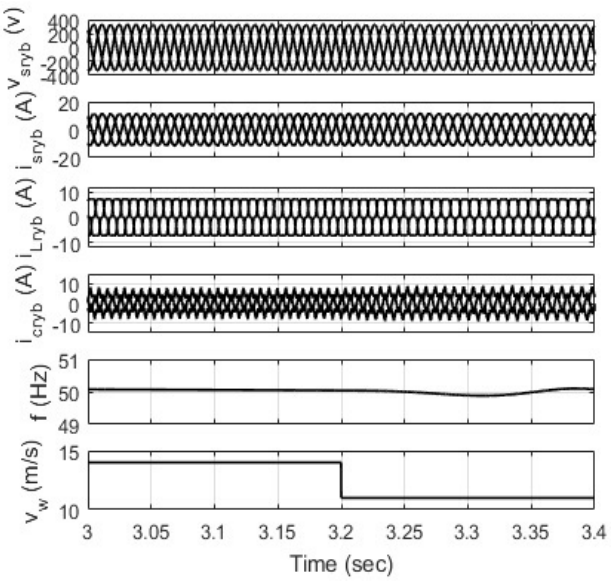
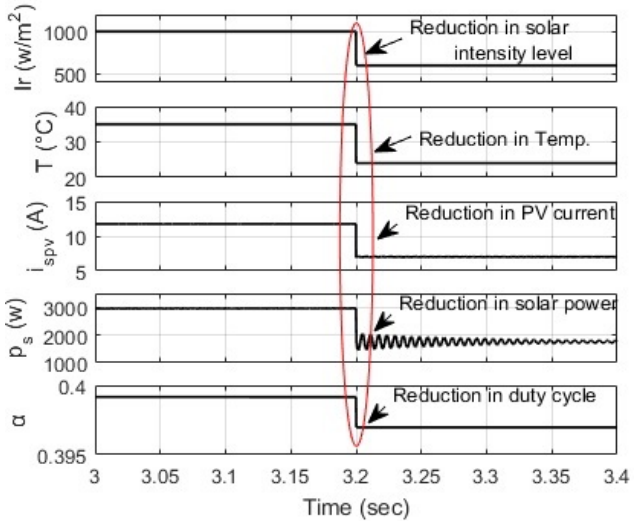


Fig. 4. Dynamic response of MG with change in wind velocity

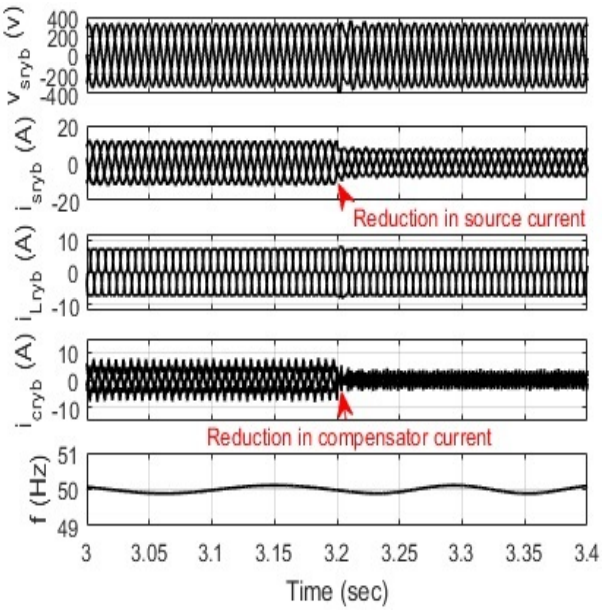
B. Dynamic response under varying solar intensity

Fig. 5(a) and (b) shows the dynamic response of isolated MG under change in solar insolation level and constant wind velocity feeding nonlinear load. Fig. 5(a) shows the change in solar intensity level and temp. (T) are from 1000 w/m^2 to 600 w/m^2 and 35 °C to 24 °C at t=3.2 s. During the variation in solar level the SPV current, duty cycle (α) are reduced. The solar power is reduced from 3kW to 1.8 kW. Fig 5(b) shows that during the change in solar level, the source current and compensator current are reduced at t=3.2 s. The DSC-FLL based controller stabilizes the frequency and source voltage of isolated MG under dynamic conditions. The source power is reduced from 5.7 kW to 3.2 kW during this change at t=3.2 s.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:25:23 UTC from IEEE Xplore. Restrictions apply.



(a)



(b)

Fig. 5(a) and (b). Dynamic response under change in solar intensity level

C. Dynamic response of MG with unbalanced load

Fig. 6 shows the dynamic response of isolated MG under fixed wind velocity and solar intensity feeding unbalanced nonlinear load. The load is disconnected for phase 'r' at t=3.1s to 3.3s. During the disconnection of load, the load current is decreased and the frequency and source voltage of isolated MG is constant. During the disconnection of load, the unused power is supplied to BSS. The source power is increased 5.6 kW to 6.4 kW due to disconnection of load.

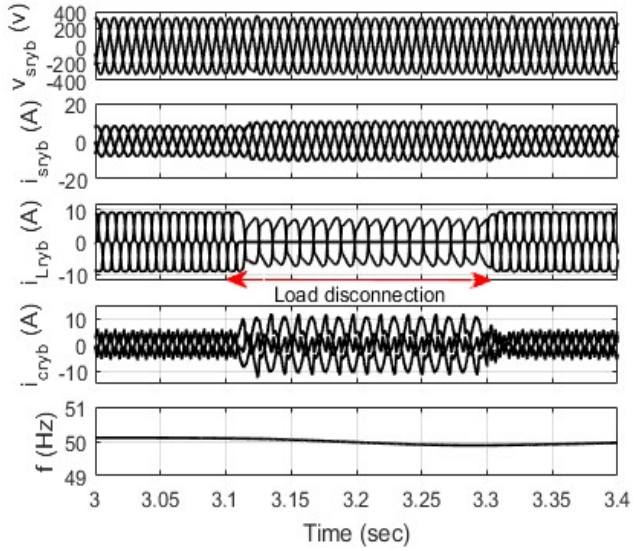


Fig. 6. Dynamic response of MG with unbalanced load

D. Battery current under different operating conditions

Fig. 7 shows the battery current various operating conditions. During the reduction in RES's (wind and solar) power, the battery current is supplied to the CCP through VSC. Hence, the BSS provides both active/reactive power supports through VSC at CCP. Moreover, the BSS consumes the extra power during minimum load demand and maintains the power between the source power and load.

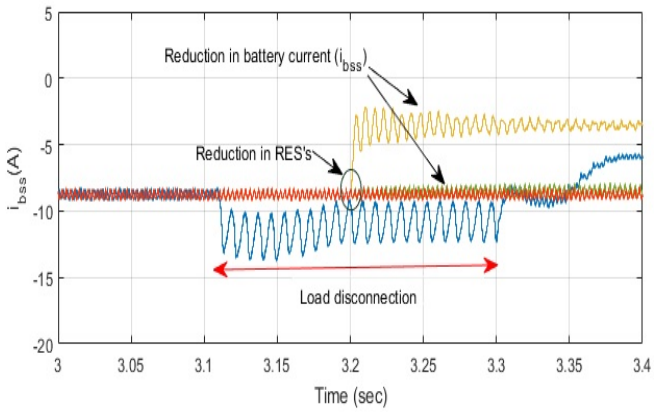
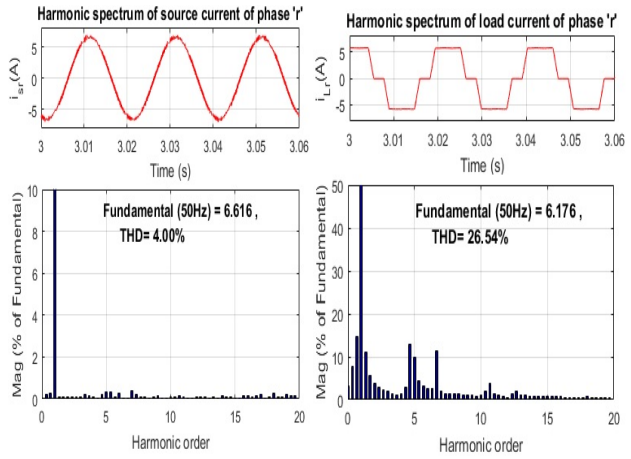


Fig..7. Battery current response under different operating conditions

E. Harmonic spectrum

Fig. 8(a) and (b) shows the harmonic spectrum of source and load current of isolated MG under nonlinear load. The THD of source current and nonlinear load current are 4 % and 26.54 % respectively. These THDs at CCP satisfy the IEEE 1547.4 standard limit. The required reactive power is provided by the controller under different dynamic conditions.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:25:23 UTC from IEEE Xplore. Restrictions apply.



(a) (b)

Fig. 8. Harmonic spectrum (a) source current and (b) nonlinear load current

V. Conclusion

The implementation of DSC-FLL and dynamic response of an isolated MG consisting of the wind driven SEIG, SPV and a BSS have been analyzed under varying load condition and variation in RES's with the DSC-FLL based control algorithm. In the presence of unbalanced and nonlinear loads, the DSC-FLL algorithm has effectively evaluated the fundamental component of load currents with reduced steady-state error and resulted in fast convergence. Further, the DSC-FLL control with VSC effectively provides the frequency and voltage stabilization at CCP, abatement of harmonics, power compensation and load balancing under intermittent condition of RES's and unbalanced loading. A modified P&O based PPPT technique has ensured maximum power transfer from the SPV. The dc-link voltage is regulated to the desired level under dynamic conditions by the BDC in the presence of BSS. The BDC also improved the battery charging current by eliminating the second harmonic component. Moreover, the CCP currents and CCP voltages have their THDs below 5% even through isolated MG is feeding nonlinear loads, hence fulfilling the requirement of the IEEE-519 standard.

References

[1] S. Sombir and M. Singh, "Voltage and Frequency Control of Self Excited Induction Generator Integrated with PV System," IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, pp. 4306-4311, 2020

[2] A. K. Giri, S. R. Arya and R. Maurya, "Compensation of Power Quality Problems in Wind-Based Renewable Energy System for Small Consumer as Isolated Loads," in IEEE Transactions on Industrial Electronics, vol. 66, no. 11, pp. 9023-9031, Nov. 2019

[3] A. K. Giri, S. R. Arya, R. Maurya and B. C. Babu, "Power Quality Improvement in Stand-Alone SEIG-Based Distributed Generation System Using Lorentzian Norm Adaptive Filter," in IEEE Transactions on Industry Applications, vol. 54, no. 5, pp. 5256-5266, Sept.-Oct. 2018

[4] B. Singh, S. S. Murthy and R. S. R. Chilipi, "STATCOM-Based Controller for a Three-Phase SEIG Feeding Single-Phase Loads," in IEEE Transactions on Energy Conversion, vol. 29, no. 2, pp. 320-331, June 2014

[5] B. Singh, R. Sharma and S. Kewat, "Robust Control Strategies for SyRG-PV and Wind-Based Islanded Microgrid," in IEEE Transactions on Industrial Electronics, vol. 68, no. 4, pp. 3137-3147, April 2021

[6] Seema Kewat, Bhim Singh and Ikhlaq Hussain, "Power Management in PV-battery-hydro based standalone microgrid," IET Renewable Power Generation. 12. 10.1049/iet-rpg.2017.0566, 2017

[7] Sombir Kundu, Madhusudan Singh and Ashutosh K. Giri, "Design and Control of a Standalone Wind–Solar System with SEIG Feeding Linear/Nonlinear Loads," In Proceedings of Symposium on Power Electronic and Renewable Energy Systems Control, LNEE, Vol 616. Springer, Singapore, 2021

[8] A. Soetedjo, I. B. Sulistiawati and Y. I. Nakhoda, "A Modified Step Size Perturb and Observe Maximum Power Point Tracking for PV System," 2019 International Conference on Engineering, Science, and Industrial Applications (ICESI), pp. 1-6, 2019

[9] P. S. Kumar, R. P. S. Chandrasena, V. Ramu, G. N. Srinivas and K. V. S. M. Babu, "Energy Management System for Small Scale Hybrid Wind Solar Battery based Microgrid," in IEEE Access, vol. 8, pp. 8336-8345, 2020

[10] Sabha Raj Arya, Sayed Alam, Jaideep Srikakolapu and Ashutosh Giri, "Phase Lock Loop–based Algorithms for DSTATCOM to Mitigate Load Created Power Quality Problems," International Transactions on Electrical Energy Systems, 2019

[11] Sepahvand Hamed Mohsen Saniei, Seyed Saeidollah Mortazavi and Saeed Golestan, "Performance Improvement of Single-Phase PLLs under Adverse Grid Conditions: An FIR Filtering-based Approach," Electric Power Systems Research, Volume 190, 106829, ISSN 0378-7796, 2021

[12] Sanchit Mishra, , Ikhlaq Hussain, Geeta Pathak, and Bhim Singh, "dPLL Based Control of A Hybrid Wind-Solar Grid Connected Inverter in the Distribution System," IET Power Electronics, 2018

[13] S. Golestan, J. M. Guerrero, J. C. Vasquez, A. M. Abusorrah and Y. Al-Turki, "A Study on Three-Phase FLLs," in IEEE Transactions on Power Electronics, vol. 34, no. 1, pp. 213-224, Jan. 2019

[14] Rohini Sharma, Seema Kewat and Bhim Singh, "Robust MMSOGI-FLL Control Algorithm for Power Quality Improvement of Solar PV-Pico Hydro-BES Based Islanded Microgrid with Dynamic Load," IET Power Electronics, 2020

[15] P. Rodríguez, A. Luna, I. Candela, R. Mujal, R. Teodorescu and F. Blaabjerg, "Multiresonant Frequency-Locked Loop for Grid Synchronization of Power Converters Under Distorted Grid Conditions," in IEEE Transactions on Industrial Electronics, vol. 58, no. 1, pp. 127-138, Jan. 2011

[16] S. Golestan, J. M. Guerrero and J. C. Vasquez, "LTP Modeling of Single-Phase T/4 Delay-Based PLLs," in IEEE Transactions on Industrial Electronics, vol. 68, no. 9, pp. 9003-9008, Sept. 2021

[17] S. Golestan, J. M. Guerrero, J. C. Vasquez, A. M. Abusorrah and Y. Al-Turki, "Linear Time-Periodic Modeling, Examination, and Performance Enhancement of Grid Synchronization Systems With DC Component Rejection/Estimation Capability," in IEEE Transactions on Power Electronics, vol. 36, no. 4, pp. 4237-4253, April 2021

[18] Ashutosh Giri, Sabha Raj Arya, Rakesh Maurya and B.C. Babu, "VCO less PLL Control Based VSC for Power Quality improvement in Distributed Generation System," IET Electric Power Applications, 2019

[19] Seema Kewat and Bhim Singh, "Modified Amplitude Adaptive Control Algorithm for Power Quality Improvement in Multiple Distributed Generation System," IET Power Electronics. 12. 2321-2329, 2019

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 04,2022 at 05:25:23 UTC from IEEE Xplore. Restrictions apply.

Opinion Leader Detection in Asian Social Networks using Modified Spider Monkey Optimization

SANJAY KUMAR, Department of Computer Science and Engineering, Delhi Technological University, India
AKSHI KUMAR*, Department of Computing and Mathematics, Manchester Metropolitan University, United Kingdom
ABHISHEK MALLIK, Department of Computer Science and Engineering, Delhi Technological University, India
SAKSHI DHALL, Department of Mathematics, Jamia Millia Islamia, India

The Asian social networks are dominated by the society's collectivist culture, and this interestingly introduces a influence mechanism aided by word-of-mouth and opinion leaders. An opinion leader can help to generate and shape other people's opinion and achieve a high information spread on any topic. In this work, a modified spider monkey optimization based opinion leader detection approach is proposed. Firstly, we employ the modified node2vec graph embedding to generate the lower dimensional vectors which act as the initial features for the nodes in a typical Asian social network. Next the entire population is broken down into several groups using the k-means++ algorithm where the number of clusters is equal to the number of opinion leaders to be selected. The local and global leaders are chosen by using the coordinates of the cluster centres of these clusters. The coordinates of the centroids of the clusters are then used to detect the local and global leaders in the network. The local leaders then form the seed set of opinion leaders for the network. The positions of the nodes in the network, including the local and global leaders, are updated over a number of iterations. At the end of these iterations, the seed set generating the maximum influence forms the set of opinion leaders in the network. We test our proposed approach using the popular information diffusion and cognitive opinion dynamics (COD) models. We perform intensive experiments on several real-life social networks based on various performance metrics. The results obtained reveal that the proposed approach outperforms several existing techniques of opinion leader detection.

CCS Concepts: • **Information systems → Web applications**; **Social networks**; • **Computing methodologies → Machine learning algorithms**.

Additional Key Words and Phrases: Asian Social Networks(ASN), Opinion Leader, Spider Monkey Optimization, k-means clustering

*Corresponding Author

Authors' addresses: Sanjay Kumar, sanjay.kumar@dtu.ac.in, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, Delhi, India, 110042; Akshi Kumar, akshi.kumar@mmu.ac.uk, Department of Computing and Mathematics, Manchester Metropolitan University, John Dalton Building, Manchester, United Kingdom; Abhishek Mallik, abhishekmallik265@gmail.com, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, Delhi, India; Sakshi Dhall, sakshidhall@gmail.com, sakshidhall@gmail.com, Department of Mathematics, Jamia Millia Islamia, New Delhi, Delhi, India.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2022 Association for Computing Machinery.
2375-4699/2022/8-ART $15.00
https://doi.org/10.1145/3555311

Just Accepted

1 INTRODUCTION

Social networking has a pulsating resplendence in the Asia-Pacific region with the highest social media penetration rates across the globe. In the era of consumerism and with the easy availability of smartphones and connected devices, the social networks in the region have witnessed a high level of engagement. These networks play an important role in generating and mobilizing the public perception regarding any topic. The collectivist behaviours among Asians add to the demographic popularity of various online networks which further stimulates social research with trust and influence leadership. There exists numerous research problems in social networks like influence maximization [25], link prediction [3, 26], community detection [23], opinion leader detection. Identifying opinion leaders is a prominent research topic in social networks analysis. Opinion leader is a crucial user with established knowledge proficiency in a particular domain whose opinions are more valued as compared to that of others. Opinion leaders are respected sources of information with sufficient interpersonal communication skills to exert influence on others' decision-making.

Formally, an opinion leader is the user having intensive knowledge in a specific domain such that he can influence the decision, behavior and opinion of other users regarding domain-related topics, ideas or products in the system [1, 5, 10, 36]. Dye [13] defined an opinion leader as a person or a set of persons that have a significant impact on customer's adoption process and decision making as compared to other users. Hence, the opinion leaders play a significant role in generating awareness and shaping public perception about any topic or issues that arise in the society. According to Gold et al. [14] opinion leader detection is a two step process. In the first step, the opinion leader analyzes, examines and understands the end-users requirements and in the second step, opinion leader generates its own opinion based on the knowledge and skills generated from the first step. Due to the large size and evolving nature of social networks, it is difficult to process such large amount of information in an efficient way for problems like finding opinion leaders. Moreover, the general representation of graph using adjacency list or adjacency matrix isn't able to capture the various features of the network in an efficient manner. Network embedding or graph embedding techniques are used to find the lower dimensional vector representations of nodes capturing their topological features in the network. This improves the computational cost of various analysis and mining task performed on these networks.

Over the years, a lot of algorithms have been proposed to optimize the solution set of the opinion leader detection problem. These algorithms aim to arrive at an optimal solution to achieve a higher quality of solution set. In the field of optimization algorithms, one of the most successful class of algorithms are the metaheuristic algorithms. A lot of research has shown that the nature-inspired metaheuristic algorithms have given excellent results for the optimization problems to achieve optimal results [47]. These algorithms develop the solution to an optimization problem by imitating the self-organising or foraging behaviour of animals or colonies appearing in the nature. The primary reason behind the effectiveness of these metaheuristic algorithms is the collaborative behaviour of these algorithms. They use their common knowledge and experience to imporve the solution. The continuous interaction amongst the various agents of the algorithms helps accumulate both the positive and negative feedback in a reward-punishment manner.

In this paper, we tackle the problem of opinion leader detection in social networks. To arrive at a solution we propose a Modified Spider Monkey Algorithm (MSMA) along with graph embedding. We modify the classical spider monkey algorithm [6] to conform with the requirements of social network analysis and suit better to the task of opinion leader detection. This algorithm is based on the foraging behaviour of the spider monkeys. We start off by first generating the modified node2vec [16] graph embedding for every node in the network. These generated embeddings act as the feature vectors for each node which is then used by our Modified Spider Monkey Algorithm. Then we initialize a population of spider monkeys with the nodes in the network being represented as the monkeys and their modified node2vec embeddings acting as the attributes of these monkeys. Then we use k-means++ algorithm to detect a global leader for the entire population. We also use k-means++ to divide the

Just Accepted

entire population into sub-groups equal to the number of the opinion leaders required. The local leader is assigned to every sub-group by using the cluster centres and these local leaders then go onto become the opinion leaders. The position of each spider monkey is updated based on the collective experience of the global leader, local leader and the monkeys belonging to its local group. The entire process is repeated until the termination condition is satisfied. The set of opinion leaders generated by our algorithm is the set of opinion leaders that generate the optimal influence spread across the various iterations of the algorithm. We investigated the validity of our proposed algorithm by conducting experiments on eight real-life large scale social networks. We also compared the performance of our algorithm against several classical algorithms and several contemporary metaheuristic based algorithms for opinion leader detection. The experiments were performed under the Independent Cascade (IC), the Susceptible-Infected-Recovered (SIR) information diffusion models. The experimental results reveal the expemplary performance of our proposed algorithm for the task of opinion leader detection by optimally identifying the critical nodes in the network. The major contributions of our work are as follows:

(i) We propose a modification of the classical spider monkey algorithm that can be utilized for the task of social network analysis.
(ii) We propose a nature-inspired metaheuristic based Modified Spider Monkey Algorithm using modified node2vec graph embedding to detect opinion leaders in a social network.
(iii) The Modified Spider Monkey Algorithm uses k-means++ to divide the population into subgroups and obtain local and global leaders for the population.
(iii) We perform exhaustive experimentation on several large scale real-life networks which reveal the commendable performance of our proposed approach for the task of opinion leader detection.

The rest of the paper is organised as follows: Section 2 discusses some of the previously done work in the field of opinion leader detection. Section 3 illustrates some of the preliminary concepts required to gain a better understanding of this work. Section 4 describes our proposed Modified Spider Monkey Algorithm in detail. The description of various real-life datasets and evaluation metrics used by us for performing the experiments is presented in Section 5. The experimental results and evaluation are presented in section 6. Section 7 concludes our work and also mentions the scope of future work.

2 RELATED WORK

The problem of opinion leader detection has been studied in a variety of ways by using differing definitions of the influence of a user in a social network. Various techniques like trust based relationships, friend-foe relationships, game theory, user status, user activities, text-mining, sentiment analysis, node centrality, ontology-based approaches and many more have been employed to detect opinion leader in a social network. Mak et al. [34] proposed a game-theory based approach for influence maximisation. They determined the null or weak associativity results for opinion-leader follower. Goyal et al. [15] presented an approach to detect opinion leaders known as the frequent pattern mining approach. As part of their approach they created a table of user action and analysed their actions. A domain specific opinion leadership approach was hypothesised by Van et al. [43] by linking opinion leaders with social network theory and anticipated that opinion leaders are usually domain specific rather than topic specific. Trusov et al. [42] presented a Bayesian shrinkage in Poisson regression using user's activity records on the premise that a user's activity influences the activities of their friends. A PageRank based approach was proposed by Heidemann et al. [18]. They proposed a quantitative approach by merging together information obtained from user's connectivity and communication activity. They tried their approach on Facebook's New Orleans dataset.

Bodendorf [8] proposed a text-mining based social network specific opinion leader detection approach. They integrated the concepts of social network analysis and text mining to unravel the opinion leaders as well as opinion trends. An ontology-based approach named BARR was propopsed by Li et al [31]. They used information

Just Accepted

like blog content, authors, readers, and their relationships hence the framework is called BARR for short. Their process is based on first creating ontology around a product and then collecting the relevant data based on BARR. Then they proceed to identify the disseminators for that data. Cho et al. [12] proposed an opinion leader detection approach based on high sociality using dispersion speed and the supreme cumulative number of adopter. An hybrid framework using a combination of novelty, activity, expertise, and influence evaluated on text contents, time and user behaviour on online learning platform was proposed by Li et al. [32]. Ma et al. [33] proposed a combination of text-mining and network topology based approach to rank super edges in the multi-dimensional model. Aghdam et al. [2] calculated the total trust value for every user in the network. Then they ranked the nodes in the network using their total trust values. Zhu et al. [48] evaluated the emotional preferences of the users using sentiment analysis. Then they created a weighted edge matrix to detect opinion leader. A dynamic social network based approach known as D_OLMiner was presented by Chen et al. [11]. They found the communities and measured the centrality of each user in a dynamic social network. njay@321 Most of the above mentioned approaches are purely algorithmic, text-mining, sentiment analysis and ontology-based approaches. Recently some centrality based and metaheuristic based approaches have also been proposed in the recent literature. Yang et al. [46] proposed a novel method based on the closeness amongst the nodes that indicated the relationships amongst various nodes of the network based on different interaction time and delays. The nodes are then ranked based on their closeness values and the nodes with the highest closeness values are considered as the opinion leaders for the networks. Jain et al. [19] proposed a fuzzy logic-based approach in which fuzzy trust rules are derived from fuzzy test systems. They also used the clustering coefficient to ascertain the prominence of the node. This was implemented together to discover the opinion leaders. An innovative social network-based nature-inspired metaheuristic based firefly algorithm was proposed by Jain et al. [20]. They discovered the opinion leaders based on the attractiveness of a node that was computed using the additive centrality of the node in the network. A whale optimization approach was also presented by Jain et al. [21]. They measured the reputation of a user in an online social network based on the various optimization functions. The opinion leader is chosen as the node with the highest reputation.

From a thorough literature survey, it is observed that very little work has been done in the area of opinion leader detection using metaheuristic based approaches. The collaborative nature of the metaheuristic algorithms and their ability to integrate the positive-feedback as well as the negative-feedback of various entities into the collective experience of the entire population makes them a viable option for opinion leader detection. This motivated us towards exploring and proposing a social network based modified spider monkey algorithm for the task of opinion leader detection.

3 PRELIMINARIES

3.1 Graph embedding

Graph embedding is used to condense a graph into a lower dimensional vector space which can be processed easily and efficiently. More formally, it can be stated that these embeddings generate a d dimensional vector for every node in the network. The vectors are oriented in such a way that the nodes which are similar to each other are placed closer together while the nodes which are dissimilar to each other are placed farther away. Some of the recently proposed network embeddings are Laplacian Eigenmaps based embeddings [7] like High-Order Proximity Preserved Embedding (HOPE) [37], Structural Deep Network Embedding (SDNE) [45], etc.

3.1.1 node2vec Graph Embedding. For our study, we have used node2vec graph embedding technique [16]. It is a random walk based embedding technique which uses random walks to better capture the local and global connectivities of the nodes in the network. This gives a clearer picture of how a node is connected in its local and global neighborhood. The node2vec graph embedding works by using a random walk from a vertex (u) and then calculating the probability of reaching another vertex (v) in that random walk. If the number of times for which the

Just Accepted

two vertices u and v appear closely in the random walks is high then these vertices are placed closely in the vector space. These generated random walks are then fed into a skip-gram neural network to generate the embeddings for a graph. It has two controlling parameters, p and q. Parameter p defines the probability of discovering the previous node while the parameter q defines the probability of discovering a previously undiscovered node or a previously undiscovered part of the graph. So basically, p controls the local exploration while q controls the global exploration. The basic idea is to make a trade-off between this local and global exploration.

The node2vec embedding method helps in optimally capturing the local and global structure details of the network using the return parameter and the in-out parameter of the random walk. The random walks also help in obtaining a generalised view of the neighborhood of the node under consideration. Moreover, the proposed modification to the classical node2vec algorithm by using the number of neighbors of the node to calculate the unnormalised transition probability gives a better sense of the dominant node in a neighborhood.

3.2 Spider Monkey Algorithm

The Spider Monkey Algorithm [6] is nature-inspired swarm-intelligence algorithm. It relies on the foraging behaviour of the spider monkey in nature. The spider monkeys are a fission–fusion social structure based animals. These animals split themselves into smaller sub-groups and merge themselves into a larger group based on the scarcity or availability of food respectively. Each sub-group has a local leader and the entire population has a global leader. Each spider monkey organises itself according to the position of local leader, global leader and monkeys from their own sub-groups. The positions of the local and global leaders are also changed based on the fitness values of the monkeys in the entire population and in their own groups. Moreover, if a local leader has not changed her position for a certain fixed number of iterations then all the monkeys of that group are redirected in a different direction. While if the global leader has not updated her position for a fixed number of times then she divides the entire group into even smaller sub-groups. If the entire population is already divided into the maximum number of sub-groups and the global leader is still not updating her position, then she combines all the sub-groups into one large group.

3.3 Degree Centrality

Degree centrality [17] assigns an importance score to a node based on the number of edges connected to the node. It works on the simple notion that the more the number of edges associated with a node, the more important it is. The degree of a node u in a network can be expressed as,

$$degree(u) = \mid N(u) \mid \tag{1}$$

here, N(u) represents the set of neighbors of node u.

3.4 Eigenvector Centrality

Eigenvector centrality [17] is used to assign an importance score to the nodes based on the notion that a node connected to high scoring nodes gets assigned a larger score than the nodes that are connected to the same number of less scoring nodes. Mathematically it can be stated as follows:

$$EC(v) = \frac{1}{\lambda} \sum_{u \in N(v), u \neq v} (A_{uv}.EC(u)) \tag{2}$$

here, N(v) represents the set of neighbors of node u. A is the adjacency matrix for the network. λ is a constant defined as follows:

$$A.EC = \lambda.EC \tag{3}$$

Just Accepted

3.5 PageRank

PageRank centrality lies at the core of google search engine has become increasing popular since its inception [9]. It builds on the notion of giving importance to nodes receiving more number of links from other nodes. Mathematically it can be expressed as,

$$pagerank(u) \;=\; \alpha \sum_{v=1}^{n} \frac{pagerank(v)}{degree_v^{out}} + \frac{1-\alpha}{n} \tag{4}$$

here, α is a damping constant similar to that in katz centrality and $degree_v^{out}$ is the outdegree for the node v.

4 PROPOSED WORK

This section illustrates our proposed work for opinion leader detection in Asian social networks(ASN) in detail. We create context-independent framework to convert the nodes of the network into lower dimensional vector space, and generate the features of the nodes in the network using modified node2vec graph embedding. Thereby, we attain a d dimensional vector representation for every node in the network. This vector represents the topological and structural details of the nodes in the network. These generated vectors then work as the feature vectors for our modified spider monkey algorithm. We modify the traditional spider monkey algorithm to optimize the objective relating to our problem statement. In this modified algorithm, the spider monkeys represent the users in the network while the local leaders form the set of the selected opinion leaders in the network. The various stages of our proposed approach are described as follows.

4.1 Feature Generation using Modified node2vec Embedding

It is often the case that due to limitations in data mining techniques, only several or sometimes none of the nodes have attributes associated with them. Moreover, even when these attributes are present, they are context-specific and thus lead to biased data analysis. Therefore, we aim to present a context-free opinion leader detection technique and for that we don't use the default attributes associated with the nodes of the network. Instead, we generate our own features based on the topological and structural details of the network. Hence, we present a modified node2vec graph embedding to generate d dimensional vector representations for the nodes. This helps to represent the nodes in a lower dimensional vector space capturing the network connectivities while remaining computationally efficient.

The proposed modified node2vec embedding generation has three phases, namely, computing transition probabilities using preprocessing, simulating random walks, and finally optimizing using Stochastic Gradient Descent (SGD). It uses return parameter, p and in-out parameter, q to guide the random walks. The return parameter (p), dictates the probability of revisiting a node immediately in the random walk. The in-out parameter (q), makes the search more localised or globalised. Let a random walk go from node u to node v and now it has to decide where to go from node v. Let the next node that the walk goes to be s. The search bias $\alpha_{pq}(u,s)$ helps to determine the unnormalised transition probability and is represented by Eq. 5.

$$\alpha_{pq}(u,s) \;=\; \begin{cases} \frac{1}{p},\; d_{us} \;=\; 0 \\ 1,\; d_{us} \;=\; 1 \\ \frac{1}{q},\; d_{us} \;=\; 2 \end{cases} \tag{5}$$

Here, d_{us} represents the shortest path between the nodes u and s. The value of d_{us} can only be either 0, 1, or 2. Using the search bias represented in Eq. 5, the unnormalised transition probability can be calculated using Eq. 6. It represents the probability of choosing a node s as the next step in the random walk. Here, w_{vs} represents the weight of the edge (v,s). For unweighted graphs, the value of w_{vs} is 1.

Just Accepted

$$\pi_{vs} = \alpha_{pq}(u,s) \cdot w_{vs} \tag{6}$$

Since, this work focuses on detecting opinion leaders and some of the network datasets available to us are unweighted, so using the weight of the edge w_{vs} isn't very relevant. Hence as part of our proposed modification, we use the total of neighbors of the node s to calculate the unnormalised transition probability. This helps in increasing the transition probability of the node which has more neighbors. It can be represented by Eq. 7. Here, $N(s)$ represents the total number of neighbors of the node s.

$$\pi_{vs} = \alpha_{pq}(u,s) \cdot N(s) \tag{7}$$

The next node in the random walks are selected based on Eq. 8. Here, c_i denotes the i^{th} node in the random walk, π_{vs} is the unnormalised transition probability in going from node v to node s, Z is the normalisation constant and is equals to the total number of transition probabilities and E is the set of all the edges. The generated random walks are then trained in the Skip-gram model in word2vec and the Stochastic Gradient Descent (SGD) algorithm and the embedding for each node is generated. Generated embeddings thereby form the d dimensional feature vectors for the nodes in the network. These generated feature vectors are then utilized by our modified spider monkey algorithm to detect opinion leaders in the network.

$$P(c_i = s \mid c_{i-1} = v) = \begin{cases} \frac{\pi_{vs}}{Z}, & (v,s) \in E \\ 0, otherwise \end{cases} \tag{8}$$

4.2 Modified Spider Monkey Algorithm

In this section, we describe our modified spider monkey algorithm. We modify the classical spider monkey algorithm to better suit the problem of opinion leader detection. Moreover, the classical spider monkey algorithm is used for continuous optimization problems whilst we work on selecting several opinion leaders given the initial population. In our proposed modification, we have modified the traditional spider monkey algorithm to better fit the opinion leader detection algorithm in social networks. We have chosen the spider monkey algorithm as it works on the principle of fission-fusion social structure amongst the animals. This divides the entire population into several sub-groups. Each sub-group has its own local leader and the entire population has a global leader as well. All the monkeys align their positions based on their local and global leaders. The concept of local leaders is utilised by us to select the opinion leaders in the sub-groups as they can better disseminate the information in their own group. This is often the case in the nature and amongst the various social networks that not every individual adheres to the opinion of the global leader but the opinion of the local leaders can still be used to motivate the majority of the population in the same direction. The global and local leaders for our network are selected using the k-means++ algorithm. The various phases of our proposed modified spider monkey algorithm are described as follows:

4.2.1 Initialization. We initialise our network dataset to conform with the concept of monkeys to fit the proposed algorithm. Hence, we treat the nodes of the network as the monkeys. We generate a d dimensional vector for every node in the network. The vectors are generated using modified node2vec graph embedding as described in section 4.1. These vectors then form the feature vectors associated with every node in the network. After the initialization process, we have n ($n =\mid V \mid$ the set of nodes in the network) spider monkeys each attributed by a d dimensional vector space and represented as follows:

$$SM_i = node2vec(i),\ i \in V \& \mid SM_i \mid = d \tag{9}$$

Where, SM_i represents the i^{th} spider monkey in the population.

Just Accepted

4.2.2 Global Leader Learning Phase. In this phase, we select a global leader for the entire network by evaluating the suitability of every node based on its feature vector. We use k-means++ clustering algorithm to detect the global leader, which helps in dividing the entire population into several sub-groups or clusters which is the crux of the spider monkey algorithm. By using k-means++ algorithm, we can map the desired number of opinion leaders to the centroids of the clusters. So by varying the value of $'k'$ in the algorithm, we can choose different number of opinion leaders (k) and predict the overall opinion spread in the network to assess the effectiveness of proposed algorithm.

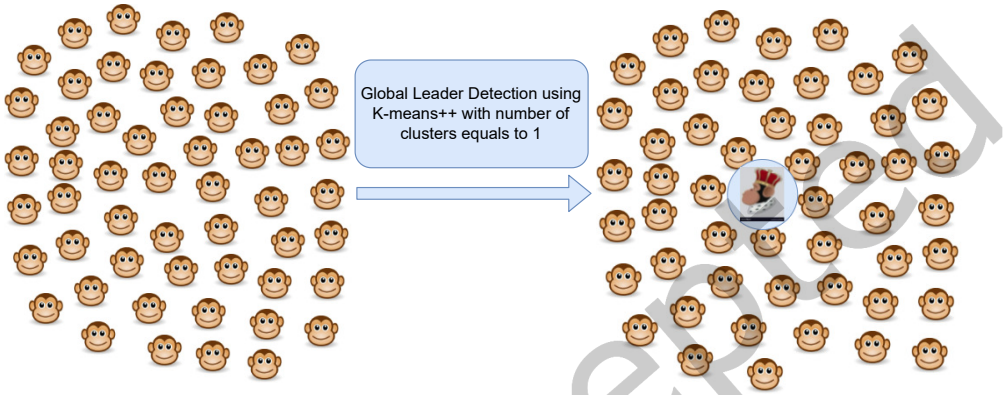


Fig. 1. Global Leader Learning Phase.

Initially, we work by assuming the entire network as a single cluster and obtain the centre of this cluster. The global leader is the spider monkey or the node closest to the centre of the aforementioned cluster. This is described in algorithm 1. Steps 1 and 2 are used to train a k-means++ model and extract the cluster centre coordinates respectively. Steps 3 to 10 generate a list, *Distances* that stores the information about the node and its distance from the cluster centre. V in step 4 represents the set of nodes in the network while d represents the chosen dimension of the graph embedding. Step 11 sorts the *Distances* list according to the distance of every node from the cluster centre in an ascending manner. Step 12 extracts the global leader finally returns it. Fig. 1 shows the procedure of selecting the global leader from a swarm of monkeys.

4.2.3 Local Leader Learning Phase. This section describes the modifications done in the local leader learning phase of the traditional spider monkey algorithm. Contrary to the classical heuristic, we don't select the local leader using a greedy approach. But instead, we use the k-means++ algorithm for selecting the local leader as we did in the global leader learning phase 4.2.2. Since the concept of local leader refers to splitting the entire population into smaller sub-groups on the basis of the similarities amongst the monkeys or nodes in the network. Therefore, we utilize the k-means++ algorithm as it splits the entire population into the desired number of clusters and also generates the cluster centres which we utilize to decide the local leaders. As the opinion leader detection problem refers to selecting k leader nodes from the entire population, so we split the entire population into k clusters and obtain the cluster centres for every cluster. The local leader of a particular cluster is the node closest to the cluster centre of that cluster. This is described in algorithm 2. Similar to algorithm 1, we first train the k-means model and extract the cluster centres in step 1 and 2. Then in step 3, we initialize a *Local_Leaders* list to store the local leaders for every sub-group in the network. Then in steps 4 to 15, we generate the local leaders for every sub-group similar to the way we generated the global leaders in algorithm 1. Finally we return a list of

Just Accepted

Algorithm 1 Algorithm for Global Leader Learning Phase

Input: Nodes of the network represented in d dimensional vector space, G
Output: Global leader
1. Model ← k-means++(G, number_of_clusters = 1)
2. Cluster ← Model.Cluster()
3. Distances := []
4. For i in range(1, V):
5. dist := 0
6. For j in range (1, d):
7. dist ← dist + sqrt(SM_{ij}^2-$Cluster_{ij}^2$)
8. end For
9. Distances.insert({i, dist})
10. end For
11. Distances ← sort_by_second(Distances)
12. Global_Leader ← Distances[0][0]
return Global_Leader

local leaders. Fig. 2 represents the process of selecting the local leaders by dividing the entire swarm into small clusters or subgroups.

Algorithm 2 Algorithm for Local Leader Learning Phase

Input: Nodes of the network represented in d dimensional vector space, G and number of desired opinion leaders, k
Output: Local leaders
1. Model ← k-means++(G, number_of_clusters = k)
2. Clusters ← Model.Clusters()
3. Local_Leaders := []
4. For g range(1, k):
5. Distances := []
6. For each member $SM_i \in g^{th}$ group:
7. dist := 0
8. For j in range (1, d):
9. dist ← dist + sqrt(SM_{ij}^2-$Cluster_{gj}^2$)
10. end For
11. Distances.insert({i, dist})
12. end For
13. Distances ← sort_by_second(Distances)
14. Local_Leaders[g] ← Distances[0][0]
15. end For
return Local_Leaders

4.2.4 Storing the Opinion Leaders. Since the major task of the opinion leaders is to simulate the maximum information diffusion across the network. So an optimal set of opinion leaders is the one which disseminates a specific opinion in an efficient and effective manner under a chosen information diffusion model (*IDM*). In

Just Accepted

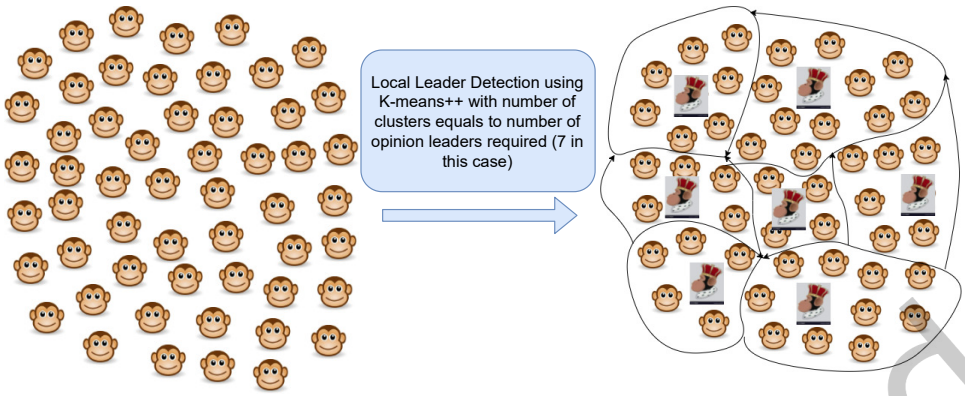


Fig. 2. Local Leader Learning Phase.

our modified spider monkey algorithm, we consider the selected local leaders to be the opinion leaders in the network. So, we run our algorithm for several iterations and for every iteration, we store the selected opinion leaders and the influence spread achieved by them. After all the iterations, we select the set of opinion leaders which performed best in terms of influence spread.

4.2.5 Local Leader Phase. This phase updates the position of every spider monkey in the d dimensional vector space using the position of the local leader and a random member of the group to which the monkey belongs to. This helps the monkeys or the users to update their positions based on the collective experience of the group. The position of the ith spider monkey belonging to the k^{th} group is updated as follows:

$$SM_{newij} = SM_{ij} + U(0,1) \times (LL_j - SM_{ij}) + U(-1,1) \times (SM_{rj} - SM_{ij}) \quad (10)$$

where SM_{ij} is the j^{th} dimension of the i^{th} spider monkey. LL_j represents the j^{th} dimension of k^{th} local group's leader. $SM_r j$ represents the $j^t h$ dimension of a random monkey chosen from the k^{th} group such that r≠i. U(l,e) is a uniformily distributed random number between l and e. Algorithm 3 illustrates the local leader phase in detail.

Algorithm 3 Algorithm for Local Leader Phase

Input: Nodes of the network represented in d dimensional vector space, G and number of desired opinion leaders, k
Output: Position updated spider monkeys, SM_{newij}

1. For g in range(1, k):
2. For each member $SM_i \in g^{th}$ group:
3. r ← random monkey in g^{th} group with r≠i
4. LL ← Local_Leaders[g]
5. For j in range(1,d):
6. $SM_{newij} = SM_{ij}+U(0,1)\times(LL_j-SM_{ij})+U(-1,1)\times(SM_{rj}-SM_{ij})$
7. end For
8. end For
9. end For

return SM_{newij}

Just Accepted

4.2.6 Global Leader Phase. In the global leader phase also, we update the position of the spider monkeys. In this phase every spider monkey updates its position based on the position of the global leader and the local members of the group, the monkey is a member of. The positions of the monkeys or users in the entire network are updated as follows.

$$SM_{newij} = SM_{ij} + U(0,1) \times (GL_j - SM_{ij}) + U(-1,1) \times (SM_{rj} - SM_{ij}) \quad (11)$$

where SM_{ij}, SM_{rj} and U(l,e) are same as described in section 4.2.5. While G_j represents the j^{th} of the global leader. This phase is further described in detail in Algorithm 4.

Algorithm 4 Algorithm for Global Leader Phase

Input: Nodes of the network represented in d dimensional vector space, G and number of desired opinion leaders, k

Output: Position updated spider monkeys, SM_{newij}

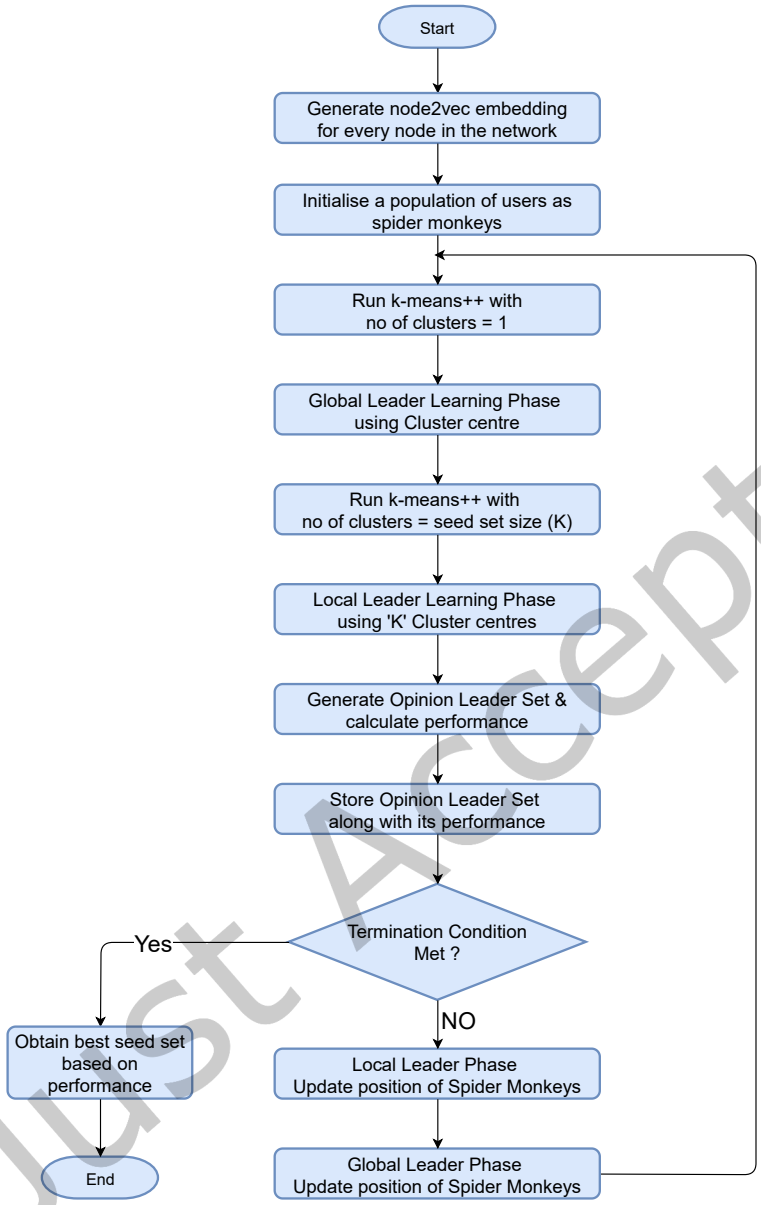
1. For g in range(1, k):
2. For each member $SM_i \in g^{th}$ group:
3. r ← random monkey in g^{th} group with r≠i
4. For j in range(1,d):
5. SM_{newij} = SM_{ij}+U(0,1)×(GL_j-SM_{ij})+U(-1,1)×(SM_{rj}-SM_{ij})
6. end For
7. end For
8. end For

return SM_{newij}

4.2.7 Termination. Any metaheuristic algorithm is associated with a termination condition. It refers to the condition which when satisfied terminates the algorithm. The termination condition is algorithm and problem specific. We also define the termination condition according to the opinion leader detection problem. Due to the lack of ground truth regarding the opinion leaders, there is no way to find out whether we obtained the best set of opinion leaders. Hence, we run our algorithm for a certain number of fixed iterations. The set of opinion leaders and their achieved performance in terms of influence spread is stored for every iteration as described in section 4.2.4. Once we have run our algorithm for all the iterations, we select the set of opinion leaders with the best performance in terms of influence spread as our set of opinion leaders.

The flow diagram and various steps of our proposed modified spider monkey algorithm is illustrated in Fig. 3. This flowchart captures the various steps described in the previous sections. It can be seen from Fig. 3 that our algorithm first generates the embedding using the modified node2vec for every node in the network. Then it goes on to initialize a population comprising of the node or users in the networks as spider monkeys with their modified node2vec embeddings as the features of the spider monkeys. Then we use k-means++ algorithms to select global leader (using 1 cluster) and local leaders (using 'K' clusters). The local leaders are then chosen as the opinion leaders and the performance of these opinion leaders in terms of influence spread is calculated and stored. Then if the termination condition is not satisfied then our algorithm updates the position of every spider monkey based on the positions of global and local leaders. It can also be seen from Fig. 3 that how our proposed approach develops on an efficient solution iteratively. It also shows that as soon as the termination condition is met, it evaluates the best solution of opinion leaders and terminates.

Just Accepted



Just Accepted

Fig. 3. Algorithmic flow chart of our proposed modified Spider Monkey Algorithm for Opinion Leader Detection.

4.3 Computational Time analysis

In this section, we estimate the computational time of our proposed Modified Spider Monkey Algorithm (MSMA). The time complexity analysis is divided into four phases: The feature generation phase, the local and global leader detection phase, the position updation phase and finally the opinion spread calculation phase. Let us consider a network $G(V, E)$ where V is the set of nodes and E is the set of edges where $\mid V \mid$ = n, and $\mid E \mid$ = m. For the first phase, we generate features using the modified node2vec embedding as shown in section 4.1. The time complexity for this phase is $O(mlogm)$ [38]. Then we go on to selecting the global and local leaders using the K-means++ algorithm as shown in section 4.2.2 and 4.2.3 respectively. The time complexity for the K-means++ algorithm turns out to be $O(nkd)$ [4], where n is the number of data points, k is the number of clusters and d is the dimensionality of the data. In our case the number of data points is the number of nodes in the network. The number of clusters, k which is 1 while detecting the global leaders and number of opinion leaders required while detecting the local leaders. The dimensionality of the data is the length of the feature vector generated by us in the first phase. The dimensionality, d and the number of clusters, k is usually a very small constant value as compared to the number of nodes in the network. Hence the time complexity for the second phase is $O(nkd) \simeq O(n)$.

Then in the third phase, we go onto updating the position of every monkey using the Local leader phase and global leader phase as shown in Section 4.2.5 and 4.2.6 respectively. In the local leader phase, we update the position of every spider monkey based on the local leader and a random member of the group to which the monkey belongs. Since we update every dimension of the position of the monkey and we do this for every monkey, so the time complexity for local leader phase is $O(nd)$. For the global leader phase, we update the position of every monkey based on the position of the global leader and a random member of the group to which the monkey belongs. This can be done with a time complexity of $O(nd)$, since here also we update every dimension of the position of every monkey. Hence the combined time complexity of the third phase is $O(nd + nd) = O(nd)$. Since d is a very small constant compared to n, so the time complexity of third phase is $O(n)$.

In the fourth phase, we calculate the influence spread of the chosen opinion leaders. Considering the case of SIR model, we find the total number of infected nodes at the end of the simulations instigated by the set of seed nodes. For the SIR model, once a node gets infected then it can infect its neighboring node with a probability of β, known as infection probability. In this study, we take β to be 0.1. To compute the spreading influence of all nodes using SIR model requires $\beta \times \sum_{v_i \in V} d_i \simeq \beta \times O(m+n) \simeq O(m+n)$ time, i.e., sum of the degree of all the nodes in a network is equal to the addition of total number of nodes and edges present in the system where d_i denotes the degree of node v_i and β is a constant. Therefore, the time complexity for the third phase is $(n + m)$. We perform the second phase and the third phase a total of $I(maximumnumberofiterations)$ times. Hence the time complexity for the combined run of second, third and fourth phase becomes $O(I(n + n + n + m))$. Since I is a small fixed constant as compare to the large number of nodes and edges, this time complexity becomes $O(I(n + n + n + m)) \simeq O(n + m)$. Hence, the total time complexity for our proposed approach is $O(mlog(m)+n+m)$. In general, the social networks are sparse networks, we can assume that $O(m) = O(n)$. Finally, the time complexity of our proposed algorithm is $O(nlogn)$.

Just Accepted

5 DATASETS AND EVALUATION METRICS

This section presents the various datasets and evaluation metrics chosen by us for our study. The datasets chosen belong to varying domains and are sampled from several online social networks. The evaluation metrics on the other hand help us to better evaluate the quality of the opinion leader set generated by our algorithm.

5.1 Dataset

This section describes the datasets chosen by us. We choose eight online social networks of varying context, size and complexities. Tab. 1 tabulates the statistical details of the networks used in our work. The description of the datasets used by us is as follows:

(i) LastFM [41]:
It represents the LastFM network of Asian users. Nodes represent the LastFM users while the edges represent the relationships between them.

(ii) Epinions [39]:
This network represents the data collected from the consumer review site Epinions.com. It is a who-trust-whom online social network.

(iii) Wiki [28, 29]:
It is the voting network representing the votes casted by users to promote an existing user to the position of administrator.

(iv) Bitcoin [24, 27]:
This is a who-trust-whom network of the users trading bitcoin on the bitcoin trading platform Bitcoin OTC.

(v) Github [40]:
It is a network of developers on an online social network for developers, Github. Nodes are developers who have starred at least ten repositories on Github and edges represent the existence of a follower-followee relationship amongst the nodes.

(vi) Deezer [41]:
This network represents the Deezer users across Europe. The nodes are the users in the network while the edges represent the follower relationships amongst the nodes.

(vii) Facebook [35]:
This network comprises of the circles or the list of friends from Facebook. The data was collected from survey participants using the Facebook app.

(viii) Email [22, 30]:
It is an email communication network from Enron. The nodes represents the users while the edges represent the existence of email communication amongst the nodes.

Just Accepted

Table 1. The basic topological features of the networks used. Here, $|V|$, $|E|$, d_{max}, d_{min}, and d_{avg} reprsents number of vertices, number of edges, maximum degrre, minimum degree, and average degree in the network, respectively.

Datasets	| **V** |	| **E** |	Type	d_{max}	d_{min}	d_{avg}
LastFM	7624	27806	Undirected	98	1	6
Epinions	75879	508837	Directed	443	1	7
Wiki	7115	103689	Directed	102	1	6
Bitcoin	5881	35592	Directed	496	1	8
Github	37700	289003	Undirected	3700	1	7
Deezer	28281	92752	Undirected	112	1	6
Facebook	4039	88234	Undirected	419	1	43
Email	36692	183831	Undirected	1400	1	10

5.2 Evaluation Metrics

This section describes the performance standards evaluated by us for ascertaining the performance prowess of our algorithm in detecting the opinion leaders. The details of these performance metrics are given below:

5.2.1 Total Opinion Spread. We evaluate the performance of our proposed model, we compare the total opinion spread of the selected opinion leaders for various algorithms. Total opinion spread can be defined as the cumulative sum of the value of opinions possessed by every node in the graph. Initially a few opinion leaders try to disseminate a particular opinion across the network. Every node in the network thereby absorbs only a fraction of that opinion depending upon the network topology. This metric is modelled for the Cognitive Opinion Dynamics model introduced by Vilone et al. [44]. The opinion of an individual can take values from 0 to 1. 1 being an alarmist opinion while 0 being a non-alarmist opinion. Initially, the opinion of the opinion leaders is modelled as 1.

5.2.2 Influence Spread. To evaluate the performance of our proposed model, we also compare the achieved influence spread of the selected opinion leaders. The selected opinion leaders form the seed set of size k which is initially activated under a desired information propagation model. Given a social network G, a seed set of opinion leaders S and an information diffusion model, then the influence spread σ(S) of the set S is the total number of seed nodes and the nodes that are activated during the spreading process under the given information diffusion model. We evaluate the influence spread achieved by our algorithm by varying several parameters as follows:

(i) Influence Spread vs. Number of Opinion Leaders, σ(S) vs. k:
It is evident that the influence spread tends to increase as the number of opinion leaders chosen increases. So we vary the size of the seed set (number of opinion leaders) to ascertain its impact on the influence spread achieved by the algorithm under consideration under different information diffusion models.

(ii) Influence Spread vs. Infection Probability:
For the IC and the SIR information diffusion models, one of the major influence control parameters is the infection probability, p and β, respectively. As the infection probability increases the influence spread also increases. Hence, we evaluate the effect of varying the infection probability, p and β for IC and SIR model on the influence spread achieved by the algorithm under consideration.

6 EXPERIMENTAL RESULTS AND ANALYSIS

This section presents and discusses the experimental setup and the experimental results obtained by us. We tested our proposed opinion leader detection method on an Asian social network, namely, LastFM and then we further generalised our approach to other social networks. In total, we have performed the experimentation on eight real-life datasets mentioned in section 5.1 and evaluated all the performance metrics mentioned in section 5.2. Intensive experiments were carried to understand the performance capabilities of our proposed approach, Modified Spider Monkey Algorithm (MSMA). The performance of our algorithm was also compared with several classical algorithms like Degree Centrality (DC) [17], Eigenvector Centrality (EC) [17], PageRank Centrality (PR) [9]. Since our proposed work is a meta heuristic based algorithm so we compare its performance with some metaheuristic based opinion leader detection techniques as well like Firefly [20] and Whale optimization [21]. All the results were obtained for the Independent Cascade (IC), the Susceptible-Infected-Recovered (SIR) As all these models are stochastic in nature, so we perform all the experiments 100 times and the obtained results are averaged out for every model. This helps to achieve more reliable results.

6.1 Total Opinion Spread vs. Number of Opinion Leaders

In this section, we compare the total opinion spread achieved by various algorithms as the number of opinion leaders are varied. We evaluate this metric for the Cognitive Opinion Dynamics (COD) based information diffusion

Just Accepted

model. We ran this model for a total of 100 iterations for every algorithm and the results of every iteration were then averaged out. Fig. 4 presents the results obtained for this metric for all the datasets and for all the chosen algorithms. As can be seen from fig. 4 that with an increase in the number of opinion leaders, the total opinion spread also increases for every network. For networks like, Epinions (Fig. 4(b)), Wiki (Fig. 4(c)), Bitcoin (Fig. 4(d)) and Github (Fig. 4(e)) the difference between the total opinion spread achieved by our algorithm and other algorithms is immense. This shows that our proposed work is a clear winner in these networks. While for the networks like Facebook (Fig. 4(a)), Deezer (Fig. 4(f)), LastFM (Fig. 4(g)) and Email (Fig. 4(h)), there tends to be a stiff competition amongst the various algorithms. But even for such networks, our algorithm performs better than other. This validates the utility of our proposed Modified Spider Monkey Algorithm (MSMA) for the task of spreading a particular opinion across the networks.

6.2 Influence Spread vs. Number of Opinion Leaders, $\sigma(S)$ vs. k

In this section, we present the influence spread attained by our proposed Modified Spider Monkey Algorithm (MSMA) and various classical and contemporary algorithms for opinion leader detection as the number of opinion leaders is varied. The results are obtained on all the real-life datasets mentioned in section 5.1 and for IC, and the SIR information diffusion model. The particular results achieved for each information diffusion model is described below:

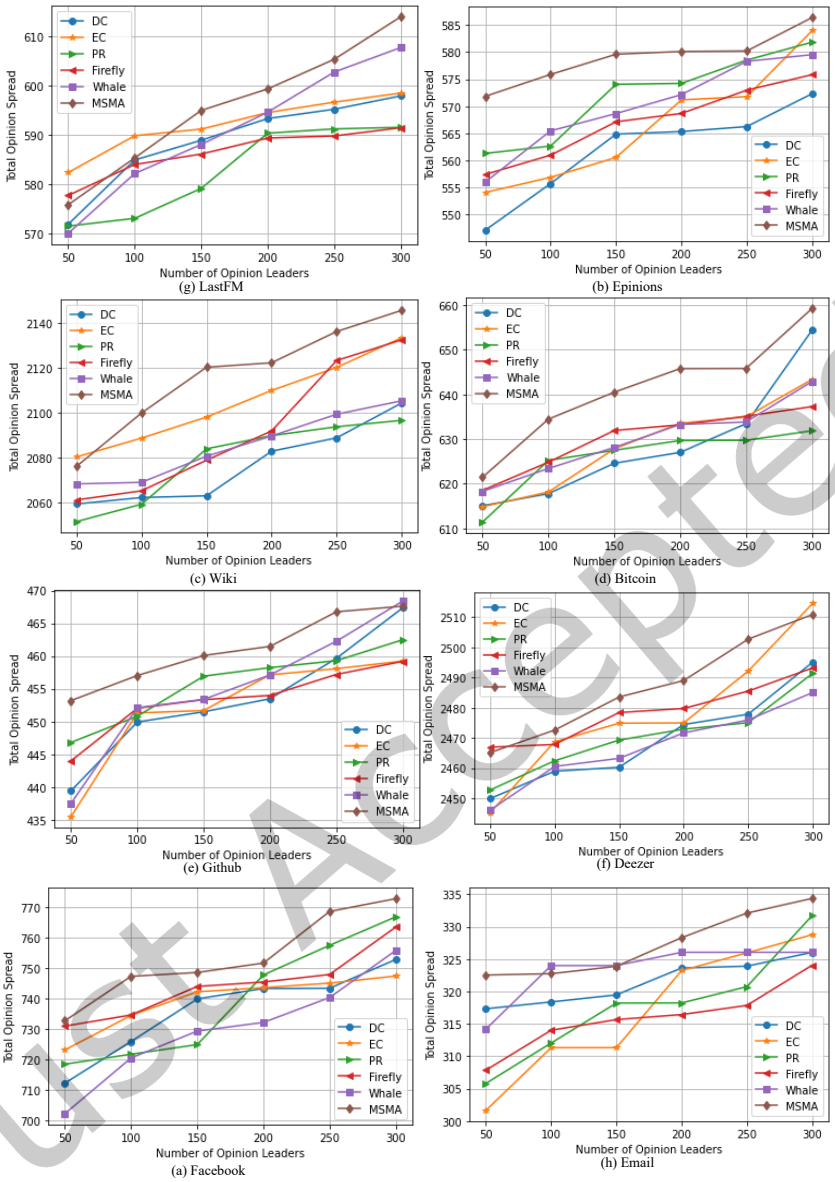
(i) Independent Cascade (IC) Model:

Fig. 5 shows the influence spread achieved by various algorithms as the number of opinion leaders is varied. The infection probability, p for the IC model is set at 0.25 for all the datasets. As can be seen from fig. 5 that the influence spread increases as the number of opinion leaders increases. This can be attributed to the fact that the as the number of initially infected population increases, the number of activated users would also during the spreading process. It can be observed that for Facebook network (fig. 5(a)), Bitcoin network (fig. 5(d)), Deezer network (fig. 5(f)) and the Email network (fig 5(h)) there is a close competition between our proposed Modified Spider Monkey Algorithm (MSMA), the whale optimization based approach and the PageRank (PR) algorithm. But still our proposed approach performs better than these algorithms and rest other algorithms as well. As for the Epinions network (fig. 5(b)), the Wiki network (fig. 5(c)), the Github network (fig. 5(e)) and the LastFM network (fig. 5(g)) our proposed algorithm clearly outperforms rest of the algorithms by a huge margin. For these networks also the Whale optimization based approach and the PageRank approach are close performers to our proposed MSMA approach. Rest of the algorithms follow thus. The above discussion shows that our proposed approach MSMA generates the best quality of opinion leaders for varying number of opinion leaders chosen.

(iii) Susceptible-Infected-Recovered (SIR) model:

Fig. 6 shows the influence spread achieved by various algorithms as the initial number of chosen opinion leaders is varied under the information diffusion model. The infection probability β is set as 0.25 for the SIR model. It can be clearly seen from fig. 6 that the proposed MSMA approach is the best performer across varying number of chosen opinion leaders for all the datasets. PageRank is the second best performer followed by whale optimization approach. The eigenvector centrality (EC) is the worst performer amongst all the compared algorithms. It can also be seen that as the number of initial opinion leaders increases the performance of our proposed algorithm becomes more and more exemplified as the difference between the achieved influence spread by our algorithm and the other algorithms increases greatly for all the datasets. This discussion corroborates the performance standards of our proposed approach.

Just Accepted

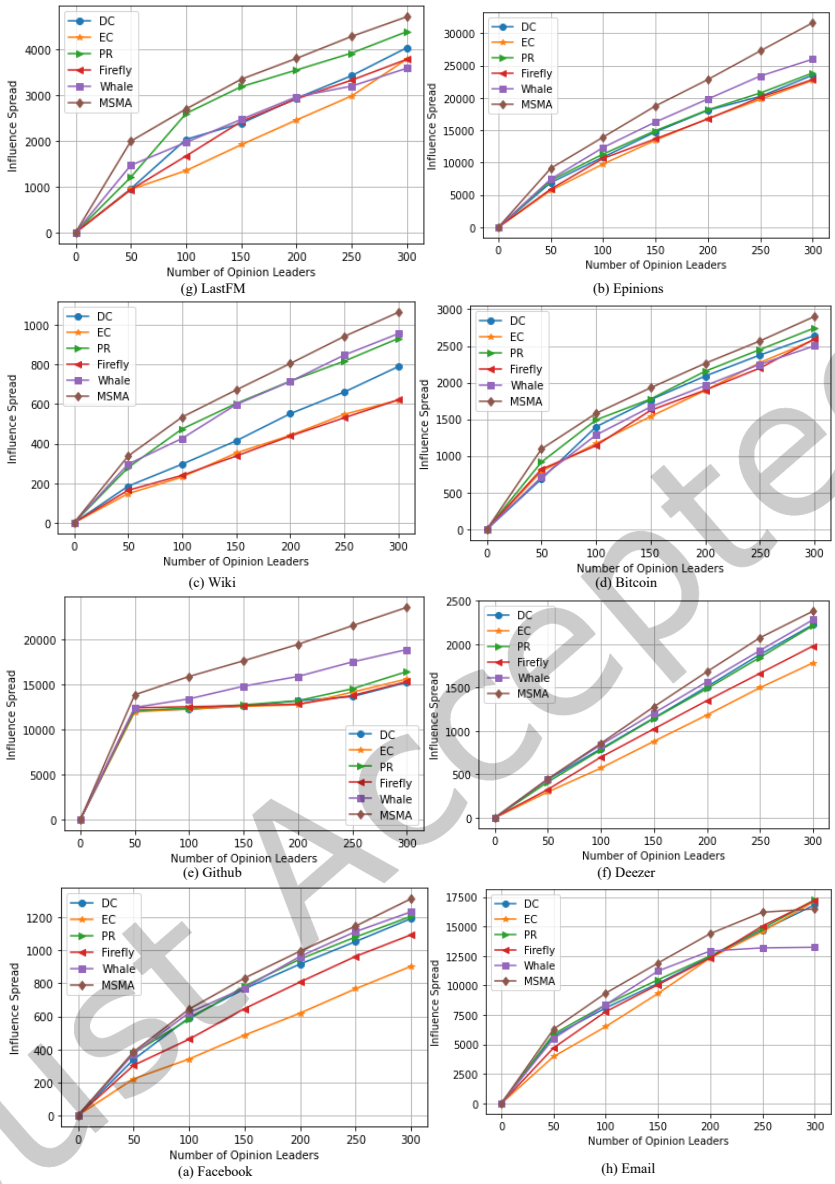


Just Accepted

Fig. 4. (a)-(h) Total opinion spread vs. number of opinion leaders value obtained by various methods on different real-life datasets. The simulations were run 100 times under the COD information diffusion model.

6.3 Influence Spread vs. Infection Probability

Now we go on to consider the effect of varying the infection probability on the influence spread achieved by various algorithms. The infection probability refers to the probability with which an activated node can infect its neighbors. The infection probability is very critical in IC and SIR information diffusion model. For the IC and the



Just Accepted

Fig. 5. (a)-(h) Final influence spread vs. number of opinion leaders value obtained by various methods on different real-life datasets. The infection probability, p is set to 0.25 under the IC model. Here, DC, EC, PR, Firefly, Whale, and MSMA refers to Degree Centrality, Eigenvector Centrality (EC), PageRank Centrality (PR), Firefly Optimization, Whale Optimization, and Modified Spider Monkey Algorithm (MSMA), respectively.

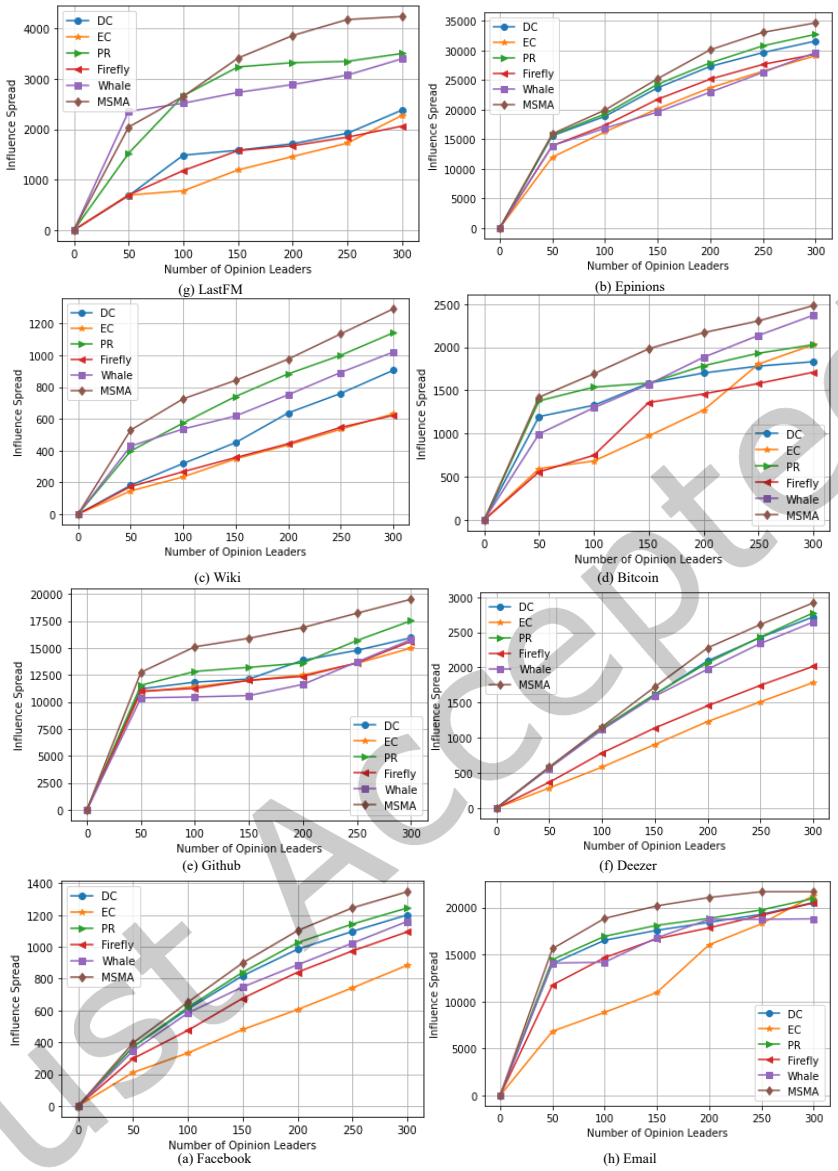


Fig. 6. (a)-(h) Final influence spread vs. number of opinion leaders value obtained by various methods on different real-life datasets. The infection probability, β is set to 0.25 under the SIR model. Here, DC, EC, PR, Firefly, Whale, and MSMA refers to Degree Centrality, Eigenvector Centrality (EC), PageRank Centrality (PR), Firefly Optimization, Whale Optimization, and Modified Spider Monkey Algorithm (MSMA), respectively.

Just Accepted

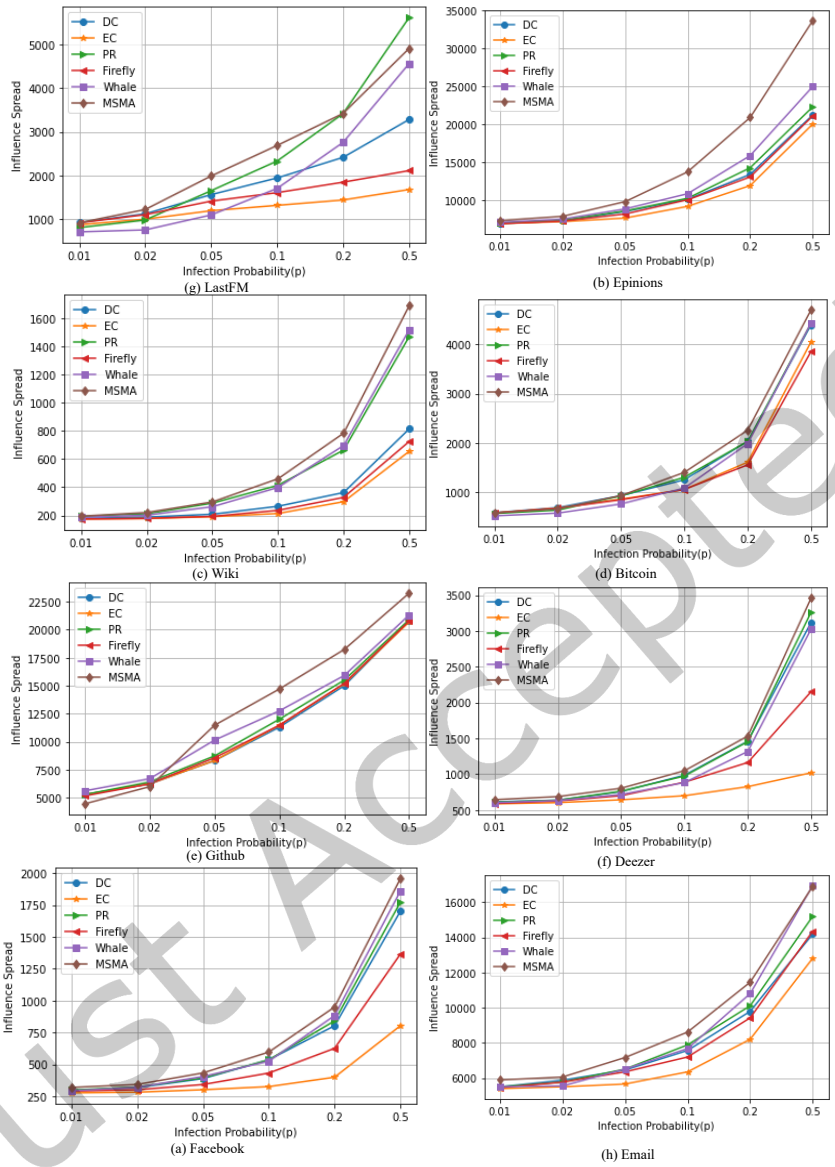
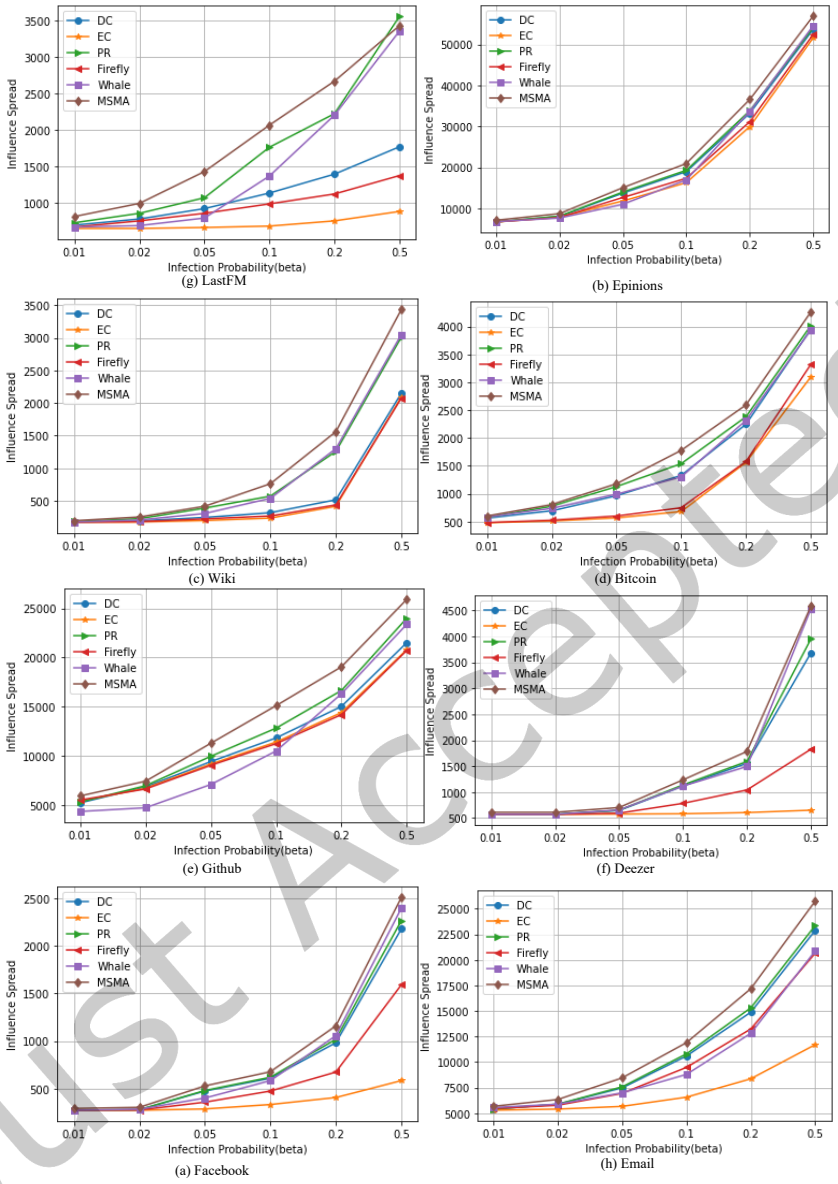


Fig. 7. (a)-(h) Final influence spread vs. infection probabilty, p value obtained by various methods on different real-life datasets. The number of opinion leaders are chosen to be 100 under the IC model. Here, DC, EC, PR, Firefly, Whale, and MSMA refers to Degree Centrality, Eigenvector Centrality (EC), PageRank Centrality (PR), Firefly Optimization, Whale Optimization, and Modified Spider Monkey Algorithm (MSMA), respectively.

Just Accepted



Just Accepted

Fig. 8. (a)-(h) Final influence spread vs. infection probabilty, β value obtained by various methods on different real-life datasets. The number of opinion leaders are chosen to be 100 under the SIR model. Here, DC, EC, PR, Firefly, Whale, and MSMA refers to Degree Centrality, Eigenvector Centrality (EC), PageRank Centrality (PR), Firefly Optimization, Whale Optimization, and Modified Spider Monkey Algorithm (MSMA), respectively.

SIR model, the higher the infection probability, the higher will be the achieved influence spread. The analysis of the effect of varying this infection probability for the IC and the SIR model is described below:

(i) Independent Cascade (IC) model:

The effect of varying the infection probability on the influence spread achieved by various algorithms under the IC model can be seen in fig. 7. The number of opinion leaders is chosen to be 100 for all the networks. It can be interpreted from the fig. 7 that the influence spread increases with an increase in infection probability for the IC model. This can be attributed to the fact that with an increase in infection probability, the ability of active nodes to activate their neighboring nodes also increases. It can be seen that for all the networks, the proposed Modified Spider Monkey Algorithm (MSMA) is the best performer. For the Epinions network (Fig. 7(b)), the Wiki network (fig. 7(c)) and the Github network (fig. 7(e)) our proposed approach gives exemplary performance and the influence spread achieved is much better than any algorithm. For the LastFM network (fig. 7(g)), our algorithms performs best until the infection probability is increased to 0.2 but as soon as the infection probability reaches 0.5 then the PageRank algorithm is the best performer. Throughout all the networks, eigenvector centrality (EC) is the worst performer. In almost all of the networks our algorithm performs best followed by whale optimisation based approach and the PageRank approach.

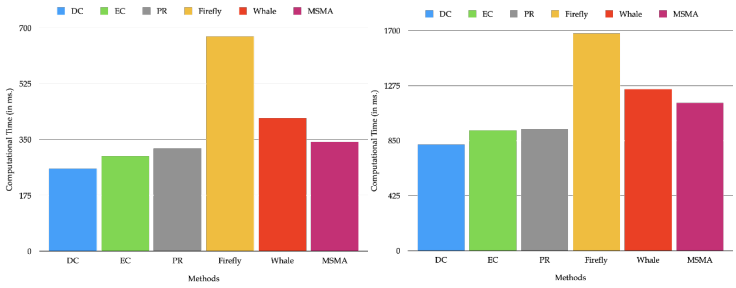
(iii) Susceptible-Infected-Recovered (SIR) model:

The effect of varying the infection probabolity (β) on the influence spread achieved by various algorithms over all the datasets under the SIR information diffusion model can be visualised in fig. 8. The number of chosen opinion leaders for the experiments were considered to be 100. It can be clearly seen from fig. 8 that as the infection probability (β) increases, the amount of influence spread achieved by various algorithms also increases. It can be observed that for the Facebook network (fig. 8(a)), Epinions network (fig. 8(b)) and the Deezer network (fig. 8(f)) the influence spread achieved by all the algorithms is very close to each other. But even in such close competition, our proposed algorithm MSMA, is the best performer. For the Wiki network (fig. 8(c)), the Bitcoin network (fig. 8(d)), the Github network (fig. 8(e)), the LastFM network (fig, 8(g)) and the Email network (fig. 8(h)), our proposed algorithm clearly outperforms other algorithms by a huge margin. Our algorithm is followed by whale optimised approch and the PageRank (PR) approach in almost all the networks. The Eigenvector Centrality (EC) based approach is the worst performer throughout the networks. The above discussion highlights the exemplary performance of our proposed Modified Spider Monkey Algorithm (MSMA) in a more reliable sense.

6.4 Absolute Computational Time

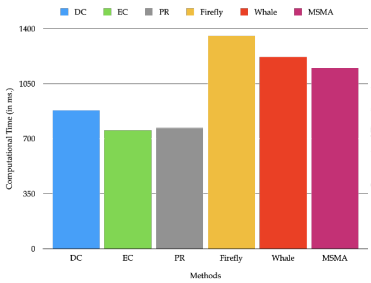
This section presents a comparison of the computational time required by various algorithms to execute. Fig. 9 represents the time taken by various algorithms to detect 100 opinion leaders for different networks. Since the metaheuristic algorithms are stochastic in nature, so we run 100 trials for every algorithm on every dataset and then average out the results. It can be clearly seen that the degree cetrality based approach for opinion leader detection takes the least amount of time. This can be explained by the fact that degree centrality based approach is a simple heuristic that considers only the degree of a node. While it can also be interpreted from the fig. 9 that the metaheuristic based approaches are the most time consuming. This can be understood by notion that the metaheuristic approaches consider a lot of parameters which are more specific to the problem under consideration and hence they take longer to execute. But even for the various metaheuristic algorithms compared by us for this study, our proposed Modified Spider Monkey Algorithm (MSMA) based approach is the best performer in terms of computational time. Moreover, the performance edge obtained by our MSMA

Just Accepted

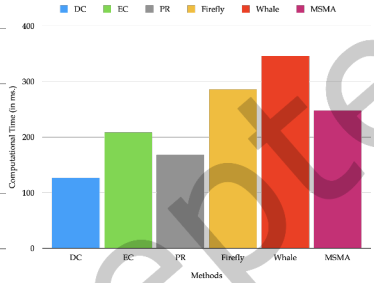


(a) Facebook

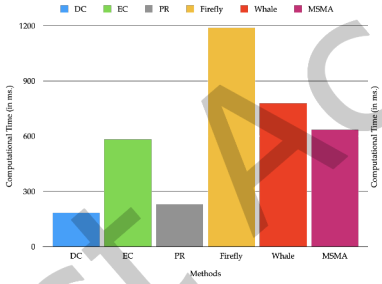
(b) Epinions



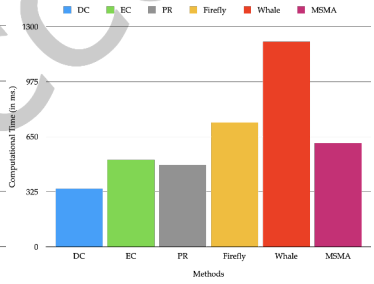
(c) Wiki



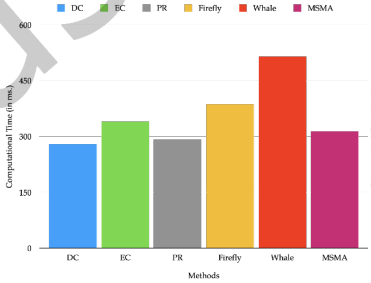
(d) Bitcoin



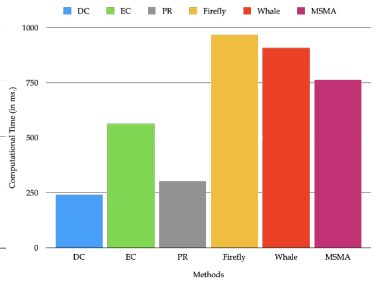
(e) Github



(f) Deezer



(g) LastFM



(h) Email

Fig. 9. (a)-(h) Comparison of the computational time required for various algorithms to select 100 opinion leaders for different datasets. The results are averaged over 100 runs.

Just Accepted

approach in terms of the influence spread achieved, more than compensates for the slightly longer run-time. The above analysis shows the plausibility of our proposed Modified Spider key Algorithm for the purpose of opinion leader detection in terms of computational time requirements as well as the influence spread targeted.

7 CONCLUSION

Social networking has gained immense popularity in the Asia-Pacific region, where billions of active users share their information and opinions. Opinion leaders tend to be the nodes in such networks which tend to have strong impact on other nodes and can dictate the opinion of a particular group in a specific direction. Through this work, we proposed an opinion detection technique in Asian social networks by leveraging the modified node2vec embedding and the spider monkey algorithm based metaheuristic. First of all, we generated the modified node2vec embedding for every node in network, this helps in generating the feature vectors for every node in the network. Then we passed the nodes along with their generated feature vectors to modified spider monkey algorithm. The classical spider monkey algorithm is modified to better fit the opinion leader detection task. We considered the nodes of the network as the spider monkeys for our study. The global and local leaders are selected by using the k-means++ algorithm instead of using the greedy approach. The chosen set of local leaders form the opinion leaders for our task. We performed intensive experimentations on several real-life social networks and compared the performance of our proposed approach by some of the classical algorithms and some of the contemporary metaheuristic algorithms for opinion leader detection. The experiments were run on Independent Cascade (IC), Susceptible-Infected-Recovered (SIR) information diffusion model. The experiments performed revealed the credibility of our proposed approach for the task of opinion leader detection as it generates a great quality of the opinion leaders. As part of the future work, we can try several different metaheuristic based approaches. We can also a hybrid mix of different metaheuristic approaches stacked together.

Just Accepted

REFERENCES

[1] Eric Abrahamson and Lori Rosenkopf. 1997. Social network effects on the extent of innovation diffusion: A computer simulation. *Organization science* 8, 3 (1997), 289–309.

[2] Samad Mohammad Aghdam and Nima Jafari Navimipour. 2016. Opinion leaders selection in the social networks based on trust relationships propagation. *Karbala International Journal of Modern Science* 2, 2 (2016), 88–97.

[3] Sameer Anand, Abhishek Mallik, Sanjay Kumar, et al. 2022. Integrating node centralities, similarity measures, and machine learning classifiers for link prediction. *Multimedia Tools and Applications* (2022), 1–29.

[4] Olivier Bachem, Mario Lucic, S Hamed Hassani, and Andreas Krause. 2016. Approximate k-means++ in sublinear time. In *Thirtieth AAAI conference on artificial intelligence*.

[5] E Bakshy, I Rosenn, C Marlow, and L Adamic. 2012. April 16-20. The role of social networks in information diffusion. *Proceedings of ACM WWW* (2012).

[6] Jagdish Chand Bansal, Harish Sharma, Shimpi Singh Jadon, and Maurice Clerc. 2014. Spider monkey optimization algorithm for numerical optimization. *Memetic computing* 6, 1 (2014), 31–47.

[7] Mikhail Belkin and Partha Niyogi. 2001. Laplacian eigenmaps and spectral techniques for embedding and clustering.. In *Nips*, Vol. 14. 585–591.

[8] Freimut Bodendorf and Carolin Kaiser. 2010. Detecting opinion leaders and trends in online communities. In *2010 Fourth International Conference on Digital Society*. IEEE, 124–129.

[9] Sergey Brin and Lawrence Page. 1998. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems* 30, 1-7 (1998), 107–117.

[10] Kenny K Chan and Shekhar Misra. 1990. Characteristics of the opinion leader: A new dimension. *Journal of advertising* 19, 3 (1990), 53–60.

[11] Yi-Cheng Chen, Lin Hui, Chun I Wu, Hsin-Yu Liu, and Sheng-Chih Chen. 2017. Opinion leaders discovery in dynamic social network. In *2017 10th International Conference on Ubi-media Computing and Workshops (Ubi-Media)*. IEEE, 1–6.

[12] Youngsang Cho, Junseok Hwang, and Daeho Lee. 2012. Identification of effective opinion leaders in the diffusion of technological innovation: A social network approach. *Technological Forecasting and Social Change* 79, 1 (2012), 97–106.

[13] Renee Dye. 2000. The buzz on buzz. *Harvard business review* 78, 6 (2000), 139–146.

[14] David Gold. 1956. Personal Influence: The Part Played by People in the Flow of Mass Communications.

[15] Amit Goyal, Francesco Bonchi, and Laks VS Lakshmanan. 2008. Discovering leaders from community actions. In *Proceedings of the 17th ACM conference on Information and knowledge management*. 499–508.
[16] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 855–864.
[17] Derek L Hansen, B Shneiderman, MA Smith, and I Himelboim. 2011. Social network analysis: measuring, mapping, and modeling collections of connections In:"Analyzing Social Media Networks with NodeXL: Insights from a Connected World".".
[18] Julia Heidemann, Mathias Klier, and Florian Probst. 2010. Identifying key users in online social networks: A pagerank based approach. (2010).
[19] Lokesh Jain and Rahul Katarya. 2018. Identification of opinion leader in online social network using fuzzy trust system. 233–239. https://doi.org/10.1109/IADCC.2018.8692095
[20] Lokesh Jain and Rahul Katarya. 2019. Discover opinion leader in online social network using firefly algorithm. *Expert systems with applications* 122 (2019), 1–15.
[21] Lokesh Jain, Rahul Katarya, and Shelly Sachdeva. 2020. Opinion leader detection using whale optimization algorithm in online social network. *Expert Systems with Applications* 142 (2020), 113016.
[22] Bryan Klimt and Yiming Yang. 2004. Introducing the Enron corpus.. In *CEAS*.
[23] Sanjay Kumar and Rahul Hanot. 2020. Community detection algorithms in complex networks: A survey. In *International Symposium on Signal Processing and Intelligent Recognition Systems*. Springer, 202–215.
[24] Srijan Kumar, Bryan Hooi, Disha Makhija, Mohit Kumar, Christos Faloutsos, and VS Subrahmanian. 2018. Rev2: Fraudulent user prediction in rating platforms. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 333–341.
[25] Sanjay Kumar, Abhishek Mallik, Anavi Khetarpal, and BS Panda. 2022. Influence maximization in social networks using graph embedding and graph neural network. *Information Sciences* (2022).
[26] Sanjay Kumar, Abhishek Mallik, and BS Panda. 2022. Link prediction in complex networks using node centrality and light gradient boosting machine. *World Wide Web* (2022), 1–27.
[27] Srijan Kumar, Francesca Spezzano, VS Subrahmanian, and Christos Faloutsos. 2016. Edge weight prediction in weighted signed networks. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*. IEEE, 221–230.
[28] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*. 641–650.
[29] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Signed networks in social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1361–1370.
[30] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. 2009. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* 6, 1 (2009), 29–123.
[31] Feng Li and Timon C Du. 2011. Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs. *Decision support systems* 51, 1 (2011), 190–197.
[32] Yanyan Li, Shaoqian Ma, Yonghe Zhang, Ronghuai Huang, et al. 2013. An improved mix framework for opinion leader identification in online learning communities. *Knowledge-Based Systems* 43 (2013), 43–51.
[33] Ning Ma and Yijun Liu. 2014. SuperedgeRank algorithm and its application in identifying opinion leader of online public opinion supernetwork. *Expert Systems with Applications* 41, 4 (2014), 1357–1368.
[34] Vincent Mak. 2008. The emergence of opinion leaders in social networks. *Available at SSRN 1157285* (2008).
[35] Julian J McAuley and Jure Leskovec. 2012. Learning to discover social circles in ego networks.. In *NIPS*, Vol. 2012. Citeseer, 548–56.
[36] James H Myers and Thomas S Robertson. 1972. Dimensions of opinion leadership. *Journal of marketing research* 9, 1 (1972), 41–46.
[37] Mingdong Ou, Peng Cui, Jian Pei, Ziwei Zhang, and Wenwu Zhu. 2016. Asymmetric transitivity preserving graph embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 1105–1114.
[38] Tiago Pimentel, Adriano Veloso, and Nivio Ziviani. 2018. Fast node embeddings: Learning ego-centric representations. (2018).
[39] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. 2003. Trust management for the semantic web. In *International semantic Web conference*. Springer, 351–368.
[40] Benedek Rozemberczki, Carl Allen, and Rik Sarkar. 2021. Multi-scale attributed node embedding. *Journal of Complex Networks* 9, 2 (2021), cnab014.
[41] Benedek Rozemberczki and Rik Sarkar. 2020. Characteristic functions on graphs: Birds of a feather, from statistical descriptors to parametric models. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 1325–1334.
[42] Michael Trusov, Anand V Bodapati, and Randolph E Bucklin. 2010. Determining influential users in internet social networks. *Journal of marketing research* 47, 4 (2010), 643–658.
[43] Ralf Van der Merwe and Gene Van Heerden. 2009. Finding and utilizing opinion leaders: Social networks and the power of relationships. *South African Journal of Business Management* 40, 3 (2009), 65–76.
[44] Daniele Vilone, Francesca Giardini, Mario Paolucci, and Rosaria Conte. 2016. Reducing individuals' risk sensitiveness can promote positive and non-alarmist views about catastrophic events in an agent-based simulation. *arXiv preprint arXiv:1609.04566* (2016).

Just Accepted

[45] Daixin Wang, Peng Cui, and Wenwu Zhu. 2016. Structural deep network embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 1225–1234.
[46] Li Yang, Yafeng Qiao, Zhihong Liu, Jianfeng Ma, and Xinghua Li. 2018. Identifying opinion leader nodes in online social networks with a new closeness evaluation algorithm. *Soft Computing* 22, 2 (2018), 453–464.
[47] Xin-She Yang. 2020. *Nature-inspired optimization algorithms*. Academic Press.
[48] Maoran Zhu, Xingkai Lin, Ting Lu, and Hongwei Wang. 2016. Identification of opinion leaders in social networks based on sentiment analysis: Evidence from an automotive forum. *Advances in Computer Science Research* 58 (2016), 412–416.

Just Accepted

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

Optimising IOT Based Smart Home Systems Using Machine-Learning Algorithm

Aman Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
amandeepnha@gmail.com

Aniket Kumar Biswas
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
skylerfrost292@gmail.com

Abhishek Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
maurya0456@gmail.com

Rajesh Kumar Yadav
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rkyadav@dtu.ac.in

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835769

***Abstract*— Our Lifestyle is getting smarter every day and technology is improving, so the advent of Internet of Things (IOT) naturally touches the security systems. The Internet connects the age-old devices to the users sitting far away using smartphones, laptops, and other gadgets. People are sitting in a much more comfortable relaxing and enjoyable environment while operating their devices. Our research addresses a crucial aspect in smart home technologies: the door locking system. We have used facial recognition combined with a Bluetooth based application to increase the capabilities to further heights. Owner can monitor who can be given the entry to the home by smart locks installed on doors. A smartphone-controlled application that uses the Bluetooth technology will be used to handle the authentication purpose. Facial recognition provides a double layer of security. Here in our system if an intruder tries to enter, we want to alert the user by firing up a buzzer and sending a message to the user who would then monitor the intrusion based upon his discretion. This approach can be extended to commercial places, schools, museums, Vending machines and all other places where entry is to be restricted.**

Keywords— Smart Home Systems, Face Recognition, HTTP, Authentication

I. INTRODUCTION

With the coming of the 21st century, humankind has made several massive leaps in terms of technology. One of the greatest and perhaps the most impactful inventions that have allowed us to prosper to such an extent is undoubtedly the Internet. The ability to wirelessly send and fetch data from one corner of the world to another at a seemingly fast rate is not groundbreaking in terms of technological achievements, but it is also, what made our significantly easier. This takes us to the notion of IoT (Internet of Things), on which this study is founded. The Internet of Things (IoT) is a network of networked computing devices that may transfer data over a network without requiring human-to-human interaction. Gone are the days when one had to rely on physical interaction to complete a task; with the internet linking many gadgets and effectively allowing device-device communication, one can do everything with a simple push of a button [4-5]. With all the factors described above in mind, it seems only reasonable that IoT technology is used to improve or perhaps create a new security system from the ground up. In this study, we chose the latter option and built a smart and safe door lock system from the ground up using these principles [9]. Furthermore, we have decided to use machine learning techniques to improve security (machine learning is the construction of computers and systems capable of self-adapting based on the study of patterns in training input data). By using the principles of computer vision in facial recognition, we aim to create a security system that not only relies on standard authorization practices but also has an added layer of impenetrable security while still maintaining the basic concept of easy-to-use minimal interaction IoT system [8]. This project stemmed from the idea that in the modern-day and age standard locks, no matter how well made cannot live up to the full potential in terms of the security provided without the use of the tools that technology provides us. A standard lock and key mechanism have several problems such as the compulsory requirement of carrying a physical key, which can get lost or if fallen into wrong hands easily duplicated. Furthermore, the complete lack of tracking systems makes it so that if anyone was to enter the home, the owner will not have the faintest idea as to who that person may be. Our system aims to combat these scenarios by creating a lock system, which replaces the need for physical keys and keeps track of which person has been granted access at what time [7].

II. LITERATURE SURVEY

This segment gives an audit on the IOT shrewd entryway locking framework and human face acknowledgment methods that utilized for enhancing the entryway lock framework. The strategies we considered for face acknowledgment methods are eigenfaces (eigenfeatures), brain organizations, HOG (Histogram of Oriented Gradients). The methodologies are investigated as far as the facial portrayals they utilized. A. IOT brilliant entryway locking framework As per the review, there are numerous frameworks as of now exist which is utilized to control the entryway and every framework is having an interesting element like framework displayed in which utilizes a LDR sensor, Ultrasonic sensor, servo engine, and laser module associated with the Arduino and Bluetooth application to foster an entryway security framework. Here, entryway can be controlled through an application utilizing Bluetooth module and gets and processes information, for example,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

power at a specific spot through Arduino UNO and utilized data set framework to store the record of individuals who have opened the entryway at what date and time. The framework displayed in mostly center around LDR sensor and ultrasonic. It doesn't have interruption location which is one of the primary downsides of framework.

B. Facial Recognition algorithms:

Eigenfaces By combining both we can drastically improve both ease of access and the Robustness of a system On the off chance that we discuss the methodologies for face acknowledgment, One of the most completely investigated approaches is eigenface. Karhunen-Loève development, eigenpicture, eigenvector, and head part are different names for it. For the portrayal of the photos of faces References utilized head part investigation. They asserted that by utilizing little assortment of loads of each face and eigenpicture (a standard face picture) could around recreated any face pictures. For acquiring the loads of each face which depict each face, project the face picture onto the eigenpicture. For the face discovery and recognizable proof References utilized eigenfaces, which was motivated by Kirby and Sirovich strategy. on each other, On the off chance that we characterize the eigenfaces as far as numerical, eigenfaces are the foremost parts of the dispersion of appearances, or the eigenvectors of the covariance framework of the arrangement of face pictures. To arrange the portrayal of various measures of the variety, individually, among the faces, eigenvectors can be utilized.

Neural Networks: Because of its non-linearity in the organization, neural networks are extremely well known. Reference utilized a solitary layer versatile organization which contains a different organization for each put away person for the face acknowledgment and this single versatile organization is one of the fake brain organizations (ANN) strategy which is additionally called WISARD. For the fruitful face acknowledgment, the significant advance is the way in developing the brain organization. We have developed a convolutional brain network in which it contains 3 convolution layers, each followed by a maximum pooling layer alongside that we had 3 thick layers with Relu enactment capacities and in the end a SoftMax initiation layer.

HOG (Histogram of Gradients) In the event that we discuss the element descriptor, which can be utilized in AI to separate various highlights from a picture and Reference involved HOG as a component descriptor which overall utilized as information base which contain the elements removed from HOGs highlights. It chips away at the guideline of separating the information picture into numerous little square cells, each containing a decent measure of pixels then Histogram of Gradients is quite possibly the most completely investigated approach since it works out the histogram of situated angles (in light of the size and direction of the edges) for every one of the pixels of the picture regarding the encompassing pixels in its cell. When both the direction and greatness are effectively determined, histograms are produced by these two qualities for every cell of the picture, bringing about a Histogram of Gradients include vector. Classifiers: Traditional classifiers like SVM, Random forest and KNN have also been used to classify faces into the correct clases. We studied the working of these algorithms and used them in our research in parallel with the HOG, Neural Network, and EigenFaces.

III. PROPOSED SYATEM

A. IDEA

The system that we have proposed works on the principle of IoT and Machine Learning. We have found that while IoT and Machine Learning are two very successful fields with no dependency on each other, By combining both we can drastically improve both ease of access and the Robustness of a system.

B. Internet of Things(IOT)

In this part of our system, we have decided to use Bluetooth as our connectivity medium and Arduino UNO as the micro-processing unit of this project. Furthermore, we have utilised the concepts of Database management to add a secure and robust authentication and authorisation service inside our system. First, the User may choose an action among three Separate choices, namely, Sign In, Sign Up and View Attendance [6].

1) SIGN UP

This is a simple form where the user is required to enter his/her details (email, name, and password) and click on ‘Sign Up. After this, these details are sent via HTTP endpoints to ‘Signup ’where the details are inserted into the existing MySQL database.

2) SIGN IN

At this point, the user is required to fill in the email address and password that he/she had registered in the system. This password and Email ID are then sent via HTTP endpoints to the PHP file ‘Signin.php’, which is hosted at the server. The proposed system also keeps track of whether the authorized user is the owner or just another registered user. The Owner has the authority to update the verification status of the other users. Once the registered user is verified, then the application allows access to the lock/unlock screen where further analysis is done based on facial recognition to ensure further security and then allows access to the Lock/Unlock button functionalities. Furthermore, once the user decides to unlock the door, the username and the current time are sent via another HTTP endpoint and PHP into the attendance database to keep track of who entered and when. System that we have proposed works on the principle of IoT and Machine Learning [2-3].

3) CHECK ATTENDANCE

By clicking on this, the user can fetch the records of all the people who have opened the lock.

C. Machine Learning Algorithms

In our proposed system the Arduino circuit is connected to a Windows PC, which runs the facial recognition parallelly on its native system [1]. This allows the code to send name and time data to the shared server whenever a face is recognized. This data is later fetched by the android application from the server and awaits user login. If the user credentials match with the data sent by the facial recognition code, then the application sends back data to the Arduino, which then

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

sends a signal to the motor to disable the lock. We decided to compare multiple machine learning and deep learning algorithms to find a suitable algorithm for the purpose of facial recognition [10]. We compared Convolutional Neural Networks (CNN), Histogram of Gradients (HOG), Principal Component Analysis (PCA), Linear discriminant Analysis (LDA), Support Vector Machines (SVM) K-nearest Neighbors (KNN) and Random Forest classifier. On our first dataset (ORL_FACES), all the algorithms performed well and gave close to 95% accuracy, but this could have been due to fewer images which were not widely dispersed over various facial expressions and occlusions. On the second dataset which was Georgia Tech having double the images compared to the previous one, we saw a decrease in the accuracy. The algorithms gave an accuracy close to 80% but Convolutional Neural Networks beat the other algorithms after showing a promising result accuracy of 85%. We used a model of 3 Convolution layers, each followed by a max-pooling layer. Then we had three dense layers with ReLu activation functions and in the end a SoftMax activation layer [11].

D. Circuit Diagram

The circuit diagram shows the circuitry being used that connects the Bluetooth signal to a physical lock.

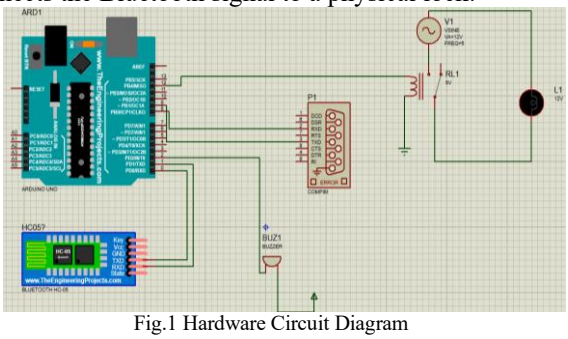


Fig.1 Hardware Circuit Diagram

E. Flow Chart

The below flow charts depicts the flow of control within the application that is responsible for controlling user access and the second one shows the flow of Bluetooth connectivity with that of facial recognition systems.

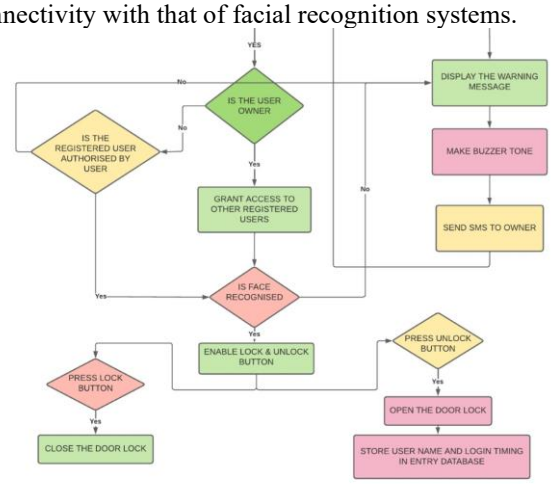


Figure 2: Flowchart

Fig.2 Access Grant flow

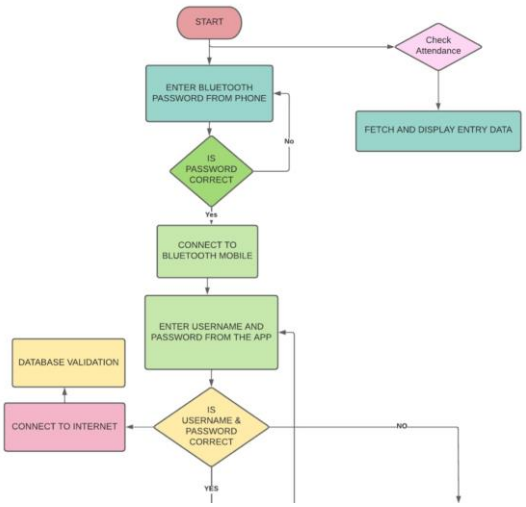


Fig.3 Bluetooth connection Flow

IV. Algorithms Compared

We decided to experiment and look at the results of 7 famous algorithms which are used extensively in the field of Facial Recognition. There were multiple research papers that compared these algorithms but we could not find a paper which compared all 7 of them together. So we decided to take the chance to examine the outcomes. In our experiment CNN performed the best. The other algorithms performed very similarly although Linear discriminant Analysis (LDA) and Histogram of Gradients (HOG) did not perform as well as the other ones.

A. Convulotianal Neural Network

Much like the human brain contains nerves and neurons that help in the transfer of audio visual data and helps us extract useful information from the rest, in machine learning too there exists Convolutional Neural Networks which work on the same principles. Each CNN model comprises mainly three separate layers, which are joined together in sequential order to form the collective model. The first layer of a Convolutional Neural Network is called a convolution layer, which is responsible for convoluting all the pixels in its receptive field into a single value. Doing this convolutes the input image and makes it ready for feature extraction. After this comes the activation layer which is responsible for applying the desired activation functions on each part of the input to determine whether a neuron should be activated or not based on the given condition. This output is then passed to the pooling layer which helps in reducing the dimensions of the resulting feature maps. This is then passed through the final layer Fully Connected Layer which transforms the entire pooled feature map into a 1D column which is then processed[13]. We used a Convolutional Neural Network model of 3 Convolution layers, each followed by a max-pooling layer. Then we had three dense layers with ReLu activation functions and in the end a SoftMax activation layer. We observed an accuracy of 95 % on the ORL_FACES dataset and an accuracy of 85% on Georgia Tech which was the best among all the algorithms.

B. Histogram of Gradients

Histogram of Gradients is a feature descriptor, which is often used in machine learning to extract different features from an image. It works on the principle of breaking the input image

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

down into multiple small square cells, each containing a fixed amount of pixels. Then it calculates the histogram of oriented gradients (based on the magnitude and orientation of the edges) for each of the pixels of the image with respect to the surrounding pixels in its cell. Once both the orientation and magnitude are successfully calculated , histograms are generated by these two values for each cell of the image, resulting in a Histogram of Gradients feature vector. The accuracy was 96 % on the ORL_FACES dataset and 69 % on the Georgia Tech dataset [15].

C. Principal Component Analysis(PCA)

PCA stands for Principal Component Analysis. It is an algorithm used for dimensionality reduction. Here in PCA we mainly reduce an n dimensional vector into some m dimensional vector where m is less than n. This is mainly done for visualisation and for reducing the number of variables so that processing can be done faster. Some accuracy might be lost because the dataset is reduced to a lower dimension but this brings an added advantage of speed which leads to faster calculation and easy to find correlations between features. In Principal Component Analysis at first the variables are standardised. After that, the covariance matrix is calculated[25]. The third step involves the calculation of eigenvectors and eigenvalues which is then followed by extracting the feature vector. Finally the data is projected onto the principal component axes. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. In our experiment Principal Component Analysis gave good results with accuracy of 96 % with the ORL_FACES dataset. With the Georgia Tech dataset the accuracy was reduced to 78%. One thing that was observed was that the best results with Principal Component Analysis were achieved with that of SVM classifier using a linear kernel.

D. Linear Discriminant Analysis(LDA)

LDA stands for Linear Discriminant Analysis is another dimensionality reduction algorithm. It separates the dataset into classes rather than just projecting them into a plane like PCA. It is due to this that the LDA can be used in classification problems directly. It does not need any classifier. It is self-sufficient in its own. In our paper we have used LDA directly to predict the classes for facial recognition. Though it can be used together with classifiers with SVM, Random Forest or KNN, etc. In our experiment we got inferior results from LDA as compared to PCA and CNN. LDA had an accuracy of 94 % with the ORL_FACES dataset. The accuracy was 70% with Georgia Tech dataset which was very less when compared with that of CNN which possessed an accuracy of 85% and even PCA with an accuracy of 78%.

E. Support Vector Analysis(SVM)

Support vector machine is a supervised learning algorithm, which is often used for classification as well as regression problems in machine learning. Support vector machine create a hyperplane with the help of support vectors which is also called as decision boundary, that help to segregate n-D(Dimensional) space into classes and in the future which help to place data point in their respective/correct category. We tried with different kernels in Support vector machine majorly, linear, poly, rbf (radial basis function) and sigmoid and then chose the best results. We observed good results with Support vector machine on its own and also when it was used as the classifier with PCA. Support vector machine gave 96% accuracy with the ORL_FACES dataset and 78% accuracy with Georgia Tech [14].

F. Random Forest

In this algorithm the majority of random records are fetched from the dataset and a decision tree is formed on each sample. Each node in the tree is a feature and based on what features are present in the image the tree is developed further until at the end, no features are left and the image is classified. Once an input image is passed through these trees, each tree returns its own classification label. The majority of these labels are then chosen as correct. Random Forest classifier gave us an accuracy of 97 % and 79 % on the ORL_FACES dataset and Georgia Tech dataset respectively [17].

G. K-Nearest Neighbour

K-Nearest Neighbors (KNN) is a supervised learning algorithm and mostly used for classification problems. In the training phase K-Nearest Neighbors only stores the dataset, doesn't learn immediately and at classification time it performs action on dataset, and it classifies the new data based on the similarity with the available categories and which category is most like available categories, putting the new data into that category. K- K-Nearest Neighbors is a lazy learner algorithm. We used K-Nearest Neighbors on both the datasets and observed an accuracy of 95 % and 80 % with ORL_Faces dataset and Georgia Tech respectively [18-19].

V. DATASETS

While choosing the dataset, we wanted to have a database that has images that show a wide range of illumination, images with different facial expressions and if possible should have occlusions as well. For in the real world our subjects could we wearing glasses, caps or hats and the lock might be placed outside so different illumination conditions might be seen by the subjects.

A. Datasets 1

The dataset 1 that we chose was "The ORL Database of Faces". It is also known as AT&T. We will be using the team ORL_FACES for this dataset.

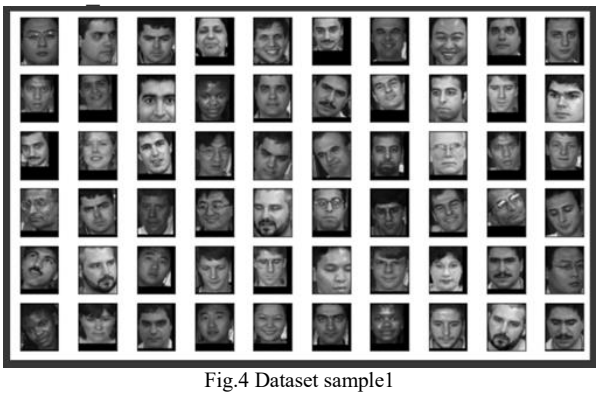


Fig.4 Dataset sample1

There are a total of 40 subjects and each subject has 10 different images available for them. The lighting in the images as well as the facial expressions and facial details

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

were varied in some of the images. The background for all these images was uniformly dark and the posture of the subjects was upright. Each image have a dimension of 92x112 pixels. The images have 256 grey levels per pixel. The background for all these images was uniformly dark and the posture of the subjects was upright.

B. Dataset 2

Fig.4 Datasets sample2

The second dataset that we chose was "Georgia Tech". This database had 750 images. Georgia Tech face database (128MB) contains images of 50 people taken in two or three sessions between 06/01/99 and 11/15/99 at the Center for Signal and Image Processing at Georgia Institute of Technology. The total number of subjects were 50 and each subject had 15 images for them. Illumination conditions were variable for each subject and the faces were captured at different orientations as well. Different facial expressions, scaling and the appearance made this dataset a good option for our research. The images in this dataset were all of JPG format but they were not of uniform size. The average size was close to 150x200 pixels; so we homogenised the images to a common size of 150x200 pixels total number of subjects were 50 and each subject had 15 images for them. Illumination conditions were variable for each subject and the faces were captured at different orientations as well.

VI. Results Analysis and Inference

We calculated the accuracy of the algorithms on both our dataset. Accuracy was the metric that we used to find our final conclusion. Since CNN had a much better accuracy in the second dataset and it was not very far behind the other ones in the first dataset, it was the algorithm we selected to have in our final algorithm. Since CNN gave an accuracy of 85 % on Georgia Tech and 95 % on ORL_FACES dataset without the use of any pre trained model, we are confident about the efficiency of the proposed model.

Algorithm	**Percentage Accuracy**	
	1.ORL_FACES	***2.Georgia Tech***
CNN	95	85
HOG	96	69
PCA	96	78
LDA	94	70
SVM	96	78
Random Forest	97	79
KNN	95	80

Table.1 Accuracy Tabulation

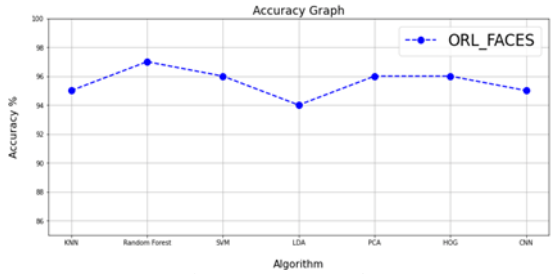


Fig.5 Accuracy Graph1

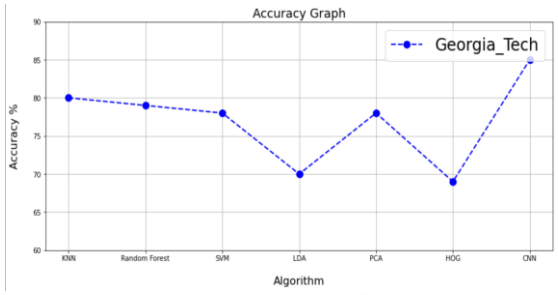


Fig.6 Accuracy Graph2

Fig 5, show the accuracy of algorithms on dataset1 and Fig 6 show the accuracy of algorithms on dataset 2.

While comparing the various algorithms, a general observation was made on the accuracy achieved on the datasets. While the ORL_Faces dataset gave high accuracy the accuracy with Georgia Tech was relatively less. Some algorithms like Principal Component Analysis, Support vector machine performed fairly well while others like Histogram of Gradients did not do well. Principal Neural Networks gave best results when it was combined with Support Vector Machines as a classifier and a linear kernel was used. Convolutional Neural Networks was the out-performer overall. Its accuracy with the Georgia Tech dataset beat the other algorithms while on the ORL_FACES dataset, the accuracy was similar to others [20]. While training Convolutional Neural Networks model we used 3 convolutional layers, each followed by a max-pooling layer[24]. Then we had three dense layers with ReLu activation functions and in the end a SoftMax activation layer. One thing that we observed was that normalisation and augmentation of data had a positive impact on the accuracy of all the algorithms improving them. Our decision to compare the results over two datasets proved to be a right choice as only the decisions over a smaller dataset like the ORL_Faces would lead to misleading results. Smart Locks can be placed outside and hence the model have to be trained over a wide variety of images with differences in lighting, occlusion and facial expressions [23].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

VII. Safety of Proposed System

The user needs to enter the username and password correctly into the mobile application and then the face must be recognized. If both the steps are successfully completed, only then the user cannot lock or unlock the door. This two-way mechanism highly improves the security of the traditional locking mechanism [22]. We have also added the buzzer and attendance record to even escalate the security system. With thefts becoming a serious issue and traditional locks being broken by physical devices, smart locks have become the need of the hour. Facial recognition is an advancement over the fingerprint technologies and combining this with records kept over the server for the username and password adds an unbreakable dual layer of security. The algorithms tested and trained ensure that the locks guarantee you security. Just facial recognition might be challenged using photographs of the concerned person, but the addition of username and password improves the locking technology thus ensuring the safety and happiness of the users [21].

VIII. Conclusion and Future Work

In this paper, we have proposed a new and innovative way to safeguard our homes. With the increase in demand for smart systems everywhere, Internet Of Things (IOT) is only going to scale upwards with no chances of going down. It is due to this reason that the locks and security systems need to advance. The Smart digital locking system combines the knowledge of software, machine learning and Internet Of Things (IOT) to improve the security. Smart locks designed as such could have a positive impact on society and our paper looks forward to it [16]. Our comparison of various machine learning and deep learning algorithms as part of our research in looking for viable facial recognition algorithms could also lead to other researches that aim similarly. Since we have got a satisfactory result with Convolutional Neural Network with using any pre trained model, we are confident of the overall efficiency of the proposed model. The face recognition in itself is difficult to beat. Combining that with the password protection of user using the application makes up for a solid locking system of the new generation homes. Our proposed system is efficient in making the security systems stronger for smart homes [20]. Analysing the faces of human beings and identifying them as early as possible is a new need of the century and our research could be beneficial to those seeking advancements in this field. With that being said our paper acts as a guide both in terms of software and hardware knowledge those who are interested in creation of appliances aimed at making homes smart.

Acknowledgment

The Authors are indebted to the central library of Delhi Technological University for providing all the access of the research material and Lab for Successfully Completing the research work. Besides it we are thankful to our Supervisor Rajesh Kumar Yadav for helping us and guiding us in each step of the research work.

References

[1] M. Shanthini, G. Vidya and R. Arun, "IoT Enhanced Smart Door Locking System," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 92-96, doi: 10.1109/ICSSIT48917.2020.9214288.

[2] Singhal, Nikita & Ganganwar, Vaishali & Yadav, Menka & Chauhan, Asha & Jakhar, Mahender & Sharma, Kareena. (2021). Comparative study of machine learning and deep learning algorithm for face recognition. Jordanian Journal of Computers and Information Technology. 07. 1. 10.5455/jjcit.71-1624859356.

[3] Vidhya, Bhojan et al. $Smart Home- Making IoT Happen." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (2020): 1281-1283.

[4] Doh, Ohsung and Ilkyu Ha. $A Digital Door Lock System for the Internet of Things with Improved Security and Usability." (2015).

[5] Satoskar, Rahul and Akarsha Mishra. $Smart Door Lock and Lighting System using Internet of Things." (2018).

[6] David, O Agbo, Madukwe Chinaza and O Odinya Jotham. $Design And Implementation Of A Door Locking System Using Android App." International Journal of Scientific & Technology Research 6 (2017): 198-203.

[7] Enhanced Security Methods of Door Locking Based Fingerprint." International Journal of Innovative Technology and Exploring Engineering (2020): n. pag.

[8] A. Özdil and M. M. Özbilen, "A survey on comparison of face recognition algorithms," 2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), 2014, pp. 1-3, doi: 10.1109/ICAICT.2014.7035956.

[9] E. Jiang, "A review of the comparative studies on traditional and intelligent face recognition methods," 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020, pp. 11-15, doi: 10.1109/CVIDL51233.2020.00010.

[10] M. Sharkas and M. A. Elenien, "Eigenfaces vs. fisherfaces vs. ICA for face recognition; a comparative study," 2008 9th International Conference on Signal Processing, 2008, pp. 914-919, doi: 10.1109/ICOSP.2008.4697276

[11] X. Fafoutis, L. Marchegiani, A. Elsts, J. Pope, R. Piechocki, and I. Craddock, "Extending the battery lifetime of wearable sensors with embedded machine learning," *IEEE World Forum Internet Things, WF-IoT 2018 - Proc.*, vol. 2018-January, pp. 269–274, 2018, doi: 10.1109/WF-IoT.2018.8355116.

[12] D. Wang, D. Chen, B. Song, N. Guizani, X. Yu, and X. Du, "From IoT to 5G I-IoT: The Next Generation IoT-Based Intelligent Algorithms and 5G Technologies," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 114–120, 2018, doi: 10.1109/mcom.2018.1701310.

[13] Q. Hu and F. Li, "Hardware design of smart home energy management system with dynamic price response," *IEEE Trans. Smart Grid*, vol. 4, no. 4,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

pp. 1878–1887, 2013, doi: 10.1109/TSG.2013.2258181.

[14] K. M. O. Nahar, B. A. Abul-Huda, A. F. Al Bataineh, and R. M. Al-Khatib, "Twins and Similar Faces Recognition Using Geometric and Photometric Features with Transfer Learning," *Int. J. Comput. Digit. Syst.*, vol. 11, no. 1, pp. 129–139, 2022, doi: 10.12785/ijcds/110110.

[15] B. K. Shah, V. Kedia, and R. K. Jha, "Integrated Vendor-Managed Time Efficient Application to Production of Inventory Systems," *Proc. 6th Int. Conf. Inven. Comput. Technol. ICICT 2021*, pp. 275–280, 2021, doi: 10.1109/ICICT50816.2021.9358504.

[16] B. K. Shah, A. Kumar, and A. Kumar, "Dog breed classifier for facial recognition using convolutional neural networks," *Proc. 3rd Int. Conf. Intell. Sustain. Syst. ICISS 2020*, pp. 508–513, 2020, doi: 10.1109/ICISS49785.2020.9315871.

[17] V. Bianchi, M. Bassoli, G. Lombardo, P. Fornacciari, M. Mordonini, and I. De Munari, "IoT Wearable Sensor and Deep Learning: An Integrated Approach for Personalized Human Activity Recognition in a Smart Home Environment," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 8553–8562, 2019, doi: 10.1109/JIOT.2019.2920283.

[18] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer internet of things devices," *Proc. - 2018 IEEE Symp. Secur. Priv. Work. SPW 2018*, no. Ml, pp. 29–35, 2018, doi: 10.1109/SPW.2018.00013.

[19] R. Shi *et al.*, "MDP and Machine Learning-Based Cost-Optimization of Dynamic Resource Allocation for Network Function Virtualization," *Proc. - 2015 IEEE Int. Conf. Serv. Comput. SCC 2015*, pp. 65–73, 2015, doi: 10.1109/SCC.2015.19.

[20] A. Shroff, B. K. Shah, A. Jha, A. K. Jaiswal, P. Sapra, and M. Kumar, "Multiplex Regulation System with Personalised Recommendation Using ML," *Proc. 5th Int. Conf. Trends Electron. Informatics, ICOEI 2021*, pp. 1567–1573, 2021, doi: 10.1109/ICOEI51242.2021.9453005.

[21] F. Jamil, N. Iqbal, Imran, S. Ahmad, and D. Kim, "Peer-to-Peer Energy Trading Mechanism Based on Blockchain and Machine Learning for Sustainable Electrical Power Supply in Smart Grid," *IEEE Access*, vol. 9, pp. 39193–39217, 2021, doi: 10.1109/ACCESS.2021.3060457.

[22] B. K. Shah, A. K. Dixit, A. K. Jaiswal, O. N. Kushwaha, A. Shroff, and N. K. Shah, "Sentiments Detection for Amazon Product Review," pp. 1–6, 2021.

[23] M. Kumar, A. Bhatia, P. Gupta, K. Jha, R. K. Jha, and B. K. Shah, "S IGN L ANGUAGE A LPHABET R ECOGNITION U SING," no. Iciccs, pp. 1859–1865, 2021.

[24] S. Akram Alrubaie, "T He R Ole of," *Rev. Lit. Arts Am.*, vol. 178, no. August, pp. 173–178, 2020.

[25] V. Kedia, A. Bhatia, S. R. Regmi, S. Dugar, K. Jha, and B. K. Shah, "Time Efficient IOS Application For CardioVascular Disease Prediction Using Machine Learning," no. Iccmc, pp. 869–874, 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:07:19 UTC from IEEE Xplore. Restrictions apply.

Hindawi Publishing Corporation
ISRN Electronics
Volume 2014, Article ID 126471, 5 pages
http://dx.doi.org/10.1155/2014/126471

Research Article

OTRA Based Voltage Mode Third Order Quadrature Oscillator

Rajeshwari Pandey,[1] Neeta Pandey,[1] Gurumurthy Komanapalli,[1] and Rashika Anurag[2]

[1] *Department of Electronics and Communication Engineering, Delhi Technological University, Delhi 110042, India*
[2] *Department of Electronics and Communication Engineering, JSS Academy of Technical Education, C-20/1 Sector 62, NOI DA, Uttar Pradesh 201301, India*

Correspondence should be addressed to Neeta Pandey; n66pandey@rediffmail.com

Received 2 February 2014; Accepted 9 March 2014; Published 29 April 2014

Academic Editors: S. Gift, R. Senani, and E. Tlelo-Cuautle

Copyright © 2014 Rajeshwari Pandey et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Two topologies of operational transresistance (OTRA) based third order quadrature oscillators (QO) are proposed in this paper. The proposed oscillators are designed using a combination of lossy and lossless integrators. The proposed topologies can be made fully integrated by implementing the resistors using matched transistors operating in linear region, which also facilitates electronic tuning of oscillation frequency. The nonideality analysis of the circuit is also given and for high frequency applications self-compensation can be used. Workability of the proposed QOs is verified through PSPICE simulations using 0.5 μm AGILENT CMOS process parameters. The total harmonic distortion (THD) for both the QO designs is found to be less than 1%.

1. Introduction

Quadrature oscillators (QO) produce outputs having a phase difference of 90°. The phase-locked sine-cosine relationship of QO has useful applications in the field of telecommunications where the modulation scheme utilizes both in-phase and quadrature components, such as single-sideband generators and quadrature mixers [1]. The QOs are also used extensively in the field of instrumentation and power electronics [2]. For these applications low value of total harmonic distortion (THD) is an essential requirement as higher harmonics have detrimental effects on electrical equipment. These higher order harmonics can also interfere with communication transmission lines since they oscillate at the same frequencies as the transmit frequency. If left unchecked, increased temperature and interference can greatly shorten the life of electronic equipment and cause damage to power systems.

It is well known that higher order networks as compared to lower order circuits provide better accuracy, frequency response, and distortion performance [3, 4]. However, it has been observed that higher order QO designs have not been explored much, as only a few third order QOs [1, 3–16] have appeared in literature in recent years. A careful observation suggests that the reported QO designs are based on forming closed loop using (i) two lossy and one lossless integrators [3–5], (ii) one lossy and two lossless integrators [8], (iii) a second order low pass filter followed by an integrator [1, 6, 7, 10–16], and (iv) three low pass filters and gained feedback around the loop [9]. These topologies are designed using second generation current conveyor (CCII) [1, 6], second generation current-controlled conveyor (CCCII) [7–9], differential voltage current conveyor (DVCC) [10], op-amp [5], operational transconductance amplifier (OTA) [3], current difference transconductance amplifier (CDTA) [4, 11], operational transresistance amplifier (OTRA) [12], and current-controlled current difference transconductance amplifier (CCCDTA) [13]. The structure proposed in [15] uses differential difference current conveyor (DDCC) and OTA whereas design of [16] is based on CCCDTA and OTA.

The OTRA is a current input voltage output device [17]. Being a current processing block, it inherits all the advantages of current mode techniques [12] and provides voltage output at low impedance which can readily be used to drive voltage input circuits without increasing component count. An extensive literature review suggests though a large number of OTRA based second order QOs [17–21] are available in literature, yet only a single third order QO topology using

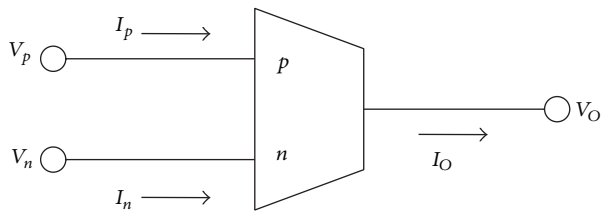


Figure 1: OTRA circuit symbol.

OTRA [12] is reported. This QO design is based on forming a closed loop using a second order low pass filter followed by an integrator.

In this paper two topologies of third order QO based on OTRA have been presented which use the scheme of lossy and lossless integrators. One of the proposed circuits makes use of one inverting lossy and two lossless integrators connected in a feedback forming closed loop whereas the other configuration uses one lossless and two lossy integrators to form closed loop. The proposed structures can be made fully integrated by implementing the resistors using matched transistors operating in linear region and can be tuned electronically.

2. Circuit Description

The OTRA is a high gain, current input voltage output analog building block [17]. The input terminals of OTRA are internally grounded, thereby eliminating response limitations due to parasitic capacitances and resistances and hence is a suitable choice for high frequency applications. The circuit symbol of OTRA is shown in Figure 1 and the port characteristics are given by (1), where R_m is transresistance gain of OTRA:

$$\begin{bmatrix} V_p \\ V_n \\ V_O \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ R_m & -R_m & 0 \end{bmatrix} \begin{bmatrix} I_p \\ I_n \\ I_O \end{bmatrix}. \quad (1)$$

For ideal operations the R_m of OTRA approaches infinity and forces the input currents to be equal. Thus OTRA must be used in a negative feedback configuration [22].

2.1. Circuit I. The first QO topology is shown in Figure 2. It uses two lossless integrators and a lossy inverting integrator in the feedback forming a closed loop resulting in loop gain of the system as $A(s)\beta(s)$, where $A(s)$ is forward path gain involving OTRA2 and OTRA3 and $\beta(s)$ is feedback gain involving OTRA1.

The criterion for oscillations [23] to occur is given by

$$1 - A(s)\beta(s) = 0. \quad (2)$$

If this criterion is satisfied, the closed loop system will result in two quadrature phase oscillations, available at nodes 1 and 2. Routine analysis of the circuit of Figure 2 results in the following characteristic equation:

$$s^3 c_1 c_2 c_3 R_2 R_3 + \frac{s^2 C_2 C_3 R_2 R_3}{R_4} + sc_3 \frac{R_3}{R_5} + \frac{1}{R_1} = 0. \quad (3)$$

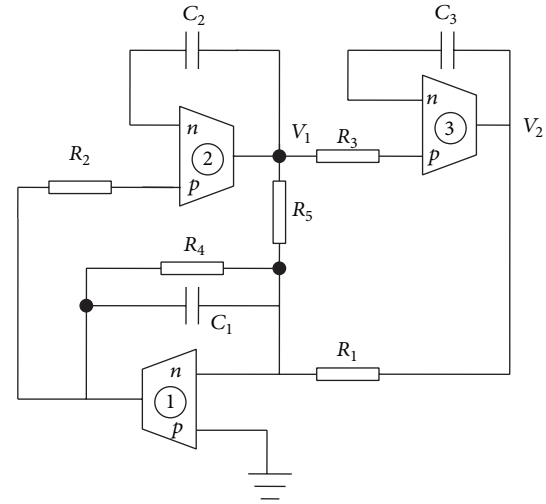


Figure 2: Proposed Circuit I.

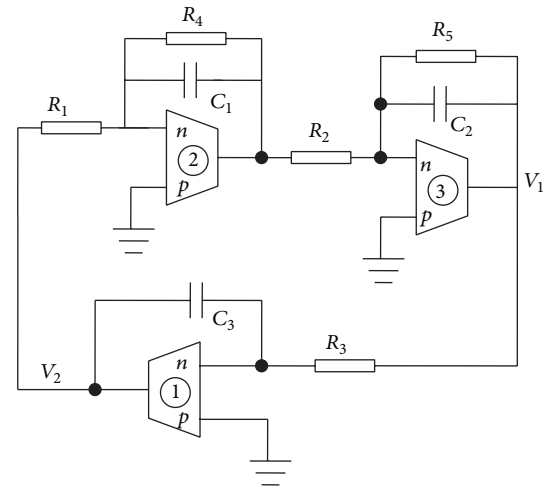


Figure 3: Proposed Circuit II.

From this characteristic equation the condition of oscillation (CO) and frequency of oscillation (FO) can be found to be FO:

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{C_1 C_2 R_2 R_5}}, \quad (4)$$

CO:

$$R_4 R_5 = R_1 R_3. \quad (5)$$

The FO can be adjusted to desired value through R_2 and proper selection of resistors R_1, R_3, and R_4 would satisfy the CO.

2.2. Circuit II. The second proposed QO configuration is shown in Figure 3 which makes use of two lossy and one lossless integrators, all in inverting mode.

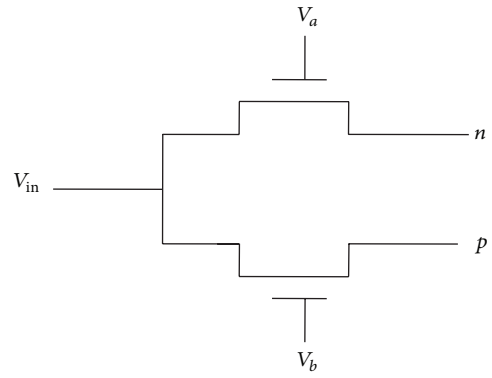


Figure 4: The MOS based resistor [17].

The characteristic equation of the Circuit II can be deduced as

$$s^3 C_1 C_2 C_3 R_3 + s^2 \left(\frac{C_2 C_3 R_3}{R_4} + \frac{C_1 C_3 R_3}{R_5} \right) + \frac{s C_3 R_3}{R_4 R_5} + \frac{1}{R_1 R_2} = 0, \tag{6}$$

FO:

$$f = \frac{1}{2\pi} \sqrt{\frac{1}{C_1 C_2 R_4 R_5}}, \tag{7}$$

CO:

$$R_1 R_2 R_3 C_3 \left[R_4 C_1 + R_5 C_2 \right] = R_4^2 R_5^2 C_1 C_2. \tag{8}$$

By suitable selection of R_4 and R_5 the FO can be adjusted to desired value and proper selection of resistors R_1, R_2, and R_3 results in desired CO.

The current differencing property of the OTRA makes it possible to implement the resistors connected to the input terminals of OTRA, using MOS transistors with complete nonlinearity cancellation [17]. Each resistor requires two matched n-MOSFETs connected in a manner as shown in Figure 4 which represents a typical MOS implementation of resistance connected at negative input of OTRA.

Symbols "p" and "n" represent the noninverting and the inverting terminals of the OTRA. As can be seen from the figure, the voltages at the drain and the source terminals for both MOSFETs are equal. On taking the difference of the currents flowing in the two transistors, the nonlinearity gets cancelled out. The resistor value realized can be expressed as

$$R = \frac{1}{K_N \left(V_a - V_b \right)}, \tag{9}$$

where

$$K_N = \mu C_{OX} \frac{W}{L}. \tag{10}$$

K_N needs to be determined for the transistors being used to implement the resistors and μ, C_{OX}, and W/L represent standard transistor parameters. The MOS based implementations of Circuit I and Circuit II are shown in Figures 5 and 6, respectively.

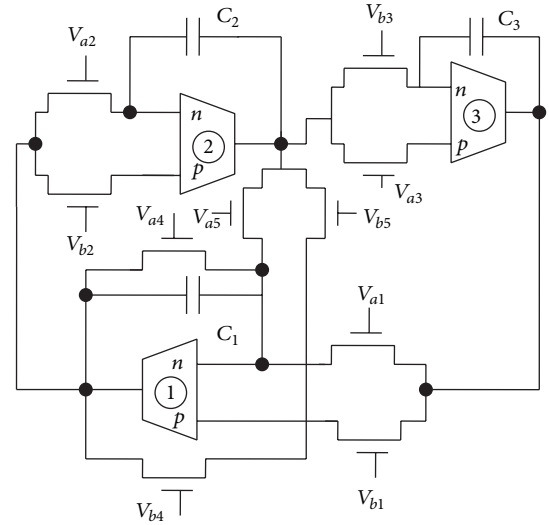


Figure 5: The MOS based implementation of QO Circuit I.

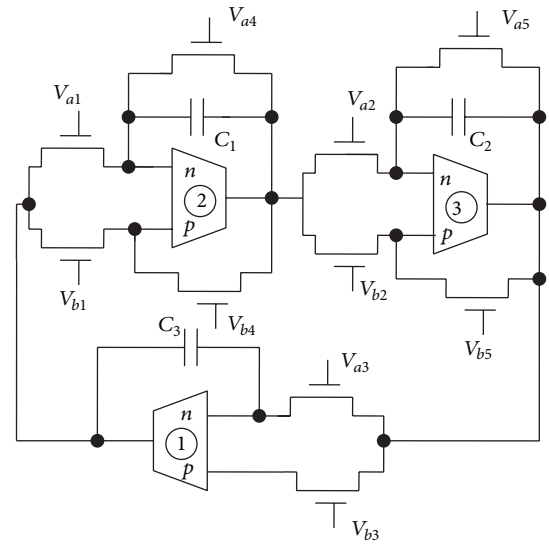


Figure 6: The MOS based implementation of QO Circuit II.

3. Nonideal Analysis

The output of the QO may deviate due to nonideality of OTRA in practice. Ideally the transresistance gain R_m is assumed to approach infinity. However, practically R_m is a frequency dependent finite value. Considering a single pole model for the transresistance gain, R_m can be expressed as

$$R_m(s) = \left(\frac{R_0}{1 + (s/\omega_0)} \right), \tag{11}$$

where R_0 is dc transresistance gain. For high frequency applications the transresistance gain $R_m(s)$ reduces to

$$R_m(s) = \left(\frac{1}{sC_p} \right), \tag{12}$$

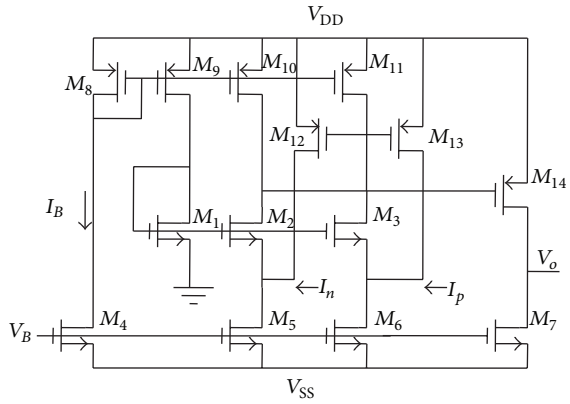


Figure 7: CMOS OTRA [22].

where

$$C_p = \frac{1}{R_0\omega_0}. \tag{13}$$

Taking this effect into account (3) modifies to

$$\begin{aligned} &s^3\left(c_1+c_p\right)\left(c_2+c_p\right)\left(c_3+c_p\right)R_2R_3 \\ &\quad + s^2\left(c_2+c_p\right)\left(c_3+c_p\right)\frac{R_2R_3}{R_4} \\ &\quad + s\left(c_3+c_p\right)\frac{R_3}{R_5}+\frac{1}{R_1}=0. \end{aligned} \tag{14}$$

From (14) it is found that FO for Circuit I changes to

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{\left(C_1+C_p\right)\left(C_2+C_p\right)R_2R_5}}. \tag{15}$$

Due to nonideality effect of OTRA (6) changes to

$$\begin{aligned} &s^3\left(c_1+c_p\right)\left(c_2+c_p\right)\left(c_3+c_p\right)R_3 \\ &\quad + s^2\left(\left(c_2+c_p\right)\left(c_3+c_p\right)\frac{R_3}{R_4}+\left(c_1+c_p\right)\left(c_3+c_p\right)\frac{R_3}{R_5}\right) \\ &\quad + s\left(c_3+c_p\right)\frac{R_3}{R_4R_5}+\frac{1}{R_1R_2}=0. \end{aligned} \tag{16}$$

The characteristic equation represented by (16) results in modified FO for Circuit II and is expressed as

$$f = \frac{1}{2\pi}\sqrt{\frac{1}{\left(C_1+C_p\right)\left(C_2+C_p\right)R_4R_5}}. \tag{17}$$

The effect of C_p can be eliminated by preadjusting the values of capacitors C_1, C_2, and C_3, thus achieving self-compensation.

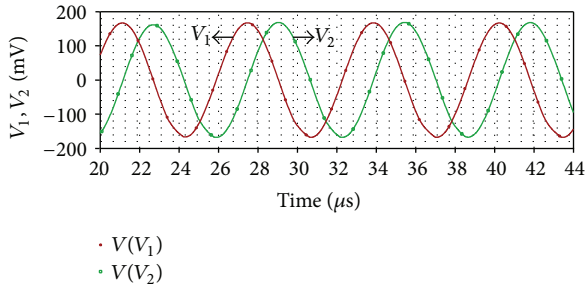


Figure 8: Transient output of proposed QO Circuit I.

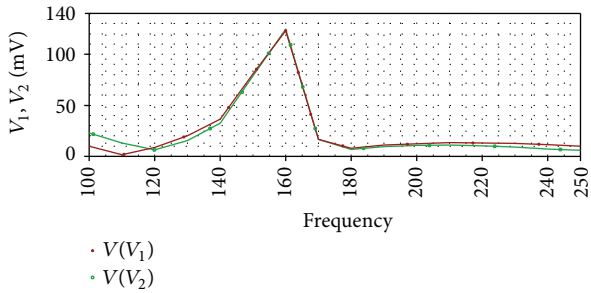


Figure 9: Frequency spectrum of output of QO Circuit I.

4. Simulation Results

The proposed QO is verified through simulations using the CMOS implementation of the OTRA [22] which is shown in Figure 7. The SPICE simulations are performed using 0.5 μm CMOS process parameters provided by MOSIS (AGILENT). Supply voltages taken are ±1.5 V. Both the QO topologies are designed for an FO of 159 KHz and the simulated value was observed to be 161 KHz for Circuits I and II, respectively. The simulated transient output and corresponding frequency spectrum for Circuit I are shown in Figures 8 and 9, respectively, and those for Circuit II are depicted in Figures 10 and 11. The percentage total harmonic distortion (THD) is 0.57% for QO Circuit I and that for Circuit II is observed to be 0.7%. These values are considerably low as compared to 6.3% THD of QO circuit of [12].

5. Conclusion

Two topologies of third order OTRA based quadrature oscillator are presented in this paper using lossy and lossless integrators. The simulations are performed using PSPICE and it is observed that the results are in close agreement with theoretical propositioned. The simulated value of % THD in both the circuits is quite low as compared to other third order structures available in literature. The effect of nonideal behavior of OTRA and high frequency compensation scheme has also been included in the proposed theory.

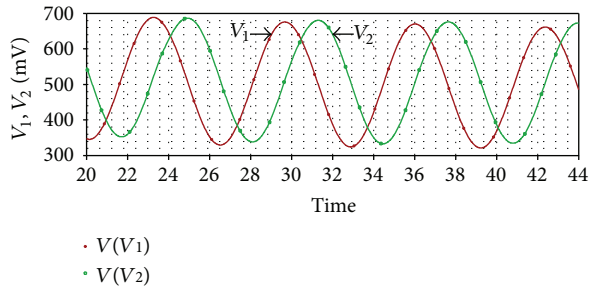


Figure 10: Transient output of proposed QO Circuit II.

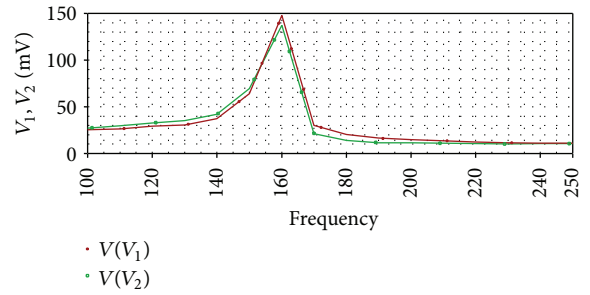


Figure 11: Frequency spectrum of output of QO Circuit II.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] J.-W. Horng, C.-L. Hou, C.-M. Chang, W.-Y. Chung, H.-W. Tang, and Y.-H. Wen, "Quadrature oscillators using CCIIs," *International Journal of Electronics*, vol. 92, no. 1, pp. 21–31, 2005.

[2] J.-W. Horng, C.-L. Hou, C.-M. Chang, H.-P. Chou, C.-T. Lin, and Y.-H. Wen, "Quadrature oscillators with grounded capacitors and resistors using FDCCIIs," *ETRI Journal*, vol. 28, no. 4, pp. 486–493, 2006.

[3] P. Prommee and K. Dejhan, "An integrable electronic-controlled quadrature sinusoidal oscillator using CMOS operational transconductance amplifier," *International Journal of Electronics*, vol. 89, no. 5, pp. 365–379, 2002.

[4] J.-W. Horng, "Current-mode third-order quadrature oscillator using CDTAs," *Active and Passive Electronic Components*, vol. 2009, Article ID 789171, 5 pages, 2009.

[5] J.-W. Horng, "Quadrature oscillators using operational amplifiers," *Active and Passive Electronic Components*, vol. 2011, Article ID 320367, 2011.

[6] J.-W. Horng, "Current/voltage-mode third order quadrature oscillator employing two multiple outputs CCIIs and grounded capacitors," *Indian Journal of Pure and Applied Physics*, vol. 49, no. 7, pp. 494–498, 2011.

[7] S. Maheshwari and I. A. Khan, "Current controlled third order quadrature oscillator," *IEE Proceedings—Circuits Devices and Systems*, vol. 152, pp. 605–607, 2005.

[8] S. Maheshwari and I. A. Khan, "Current-mode third-order quadrature oscillator," *IET Circuits, Devices and Systems*, vol. 4, no. 3, pp. 188–195, 2010.

[9] S. Maheshwari and R. Verma, "Electronically tunable sinusoidal oscillator circuit," *Active and Passive Electronic Components*, vol. 2012, Article ID 719376, 6 pages, 2012.

[10] S. Maheshwari, "Analogue signal processing applications using a new circuit topology," *IET Circuits, Devices and Systems*, vol. 3, no. 3, pp. 106–115, 2009.

[11] J.-W. Horng, H. Lee, and J.-Y. Wu, "Electronically tunable third-order quadrature oscillator using CDTAs," *Radioengineering*, vol. 19, no. 2, pp. 326–330, 2010.

[12] R. Pandey, N. Pandey, and S. K. Paul, "MOS-C third order quadrature oscillator using OTRA," in *Proceedings of the 3rd International Conference on Computer and Communication Technology*, pp. 77–80, November 2012.

[13] S. Lawanwisut and M. Siripruchyanun, "High output-impedance current-mode third-order quadrature oscillator based on CCCCTAs," in *Proceedings of the IEEE Region 10 Conference (TENCON '09)*, pp. 1–4, November 2009.

[14] A. Kwawsibsam, B. Sreewirote, and W. Jaikla, "Third-order voltage-mode quadratrue oscillator using DDCC and OTAs," in *International Conference on Circuits, System and Simulation*, pp. 317–321, 2011.

[15] M. Kumngern and S. Junnapiya, "Current-mode third-order quadrature oscillator using minimum elements," in *Proceedings of the International Conference on Electrical Engineering and Informatics (ICEEI '11)*, pp. 1–4, July 2011.

[16] D. Duangmalai and W. Jaikla, "Realization of current-mode quadrature oscillator based on third order technique," *ACEEE International Journal on Electrical and Power Engineering*, vol. 2, pp. 46–49, 2011.

[17] K. N. Salama and A. M. Soliman, "CMOS operational transresistance amplifier for analog signal processing," *Microelectronics Journal*, vol. 30, no. 3, pp. 235–245, 1999.

[18] K. N. Salama and A. M. Soliman, "Universal filters using Operational Transresistance Amplifiers," *AEU-Archiv für Elektronik und Ubertragungstechnik*, vol. 53, no. 1, pp. 49–52, 1999.

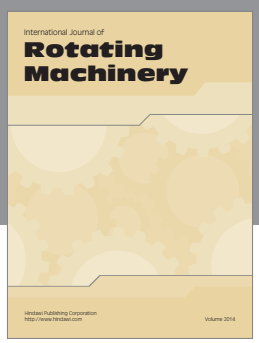
[19] S. Kilinç and U. Çam, "Operational transresistance amplifier based first-order allpass filter with an application example," in *Proceedings of the 47th Midwest Symposium on Circuits and Systems*, vol. 1, pp. 65–68, July 2004.

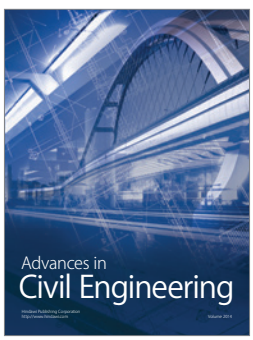
[20] S. Kilinç and U. Çam, "Cascadable allpass and notch filters employing single operational transresistance amplifier," *Computers and Electrical Engineering*, vol. 31, no. 6, pp. 391–401, 2005.

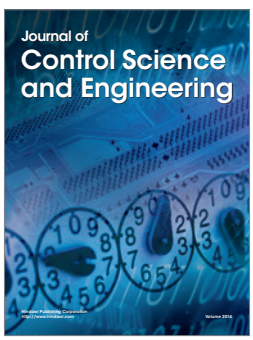
[21] C. Cakir, U. Cam, and O. Cicekoglu, "Novel allpass filter configuration employing single OTRA," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 52, no. 3, pp. 122–125, 2005.

[22] H. Mostafa and A. M. Soliman, "A modified CMOS realization of the operational transresistance amplifier (OTRA)," *Frequenz*, vol. 60, no. 3-4, pp. 70–76, 2006.

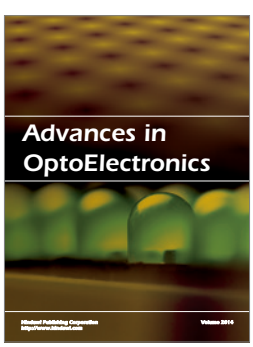
[23] A. S. Sedra and K. C. Smith, *Microelectronic Circuits.*, Oxford University Press, New York, NY, USA, 2004.

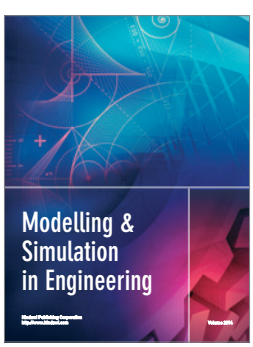


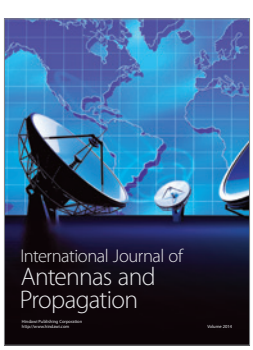


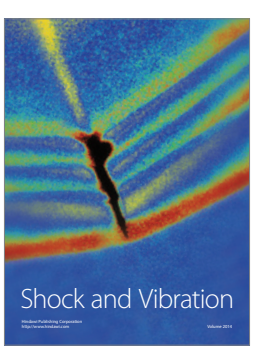


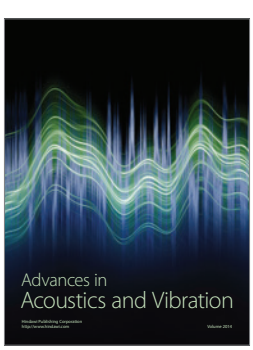
Hindawi
Submit your manuscripts at
http://www.hindawi.com











Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Parameter Optimization for Enhanced Biodiesel Yield from *Linum usitatissimum* Oil Through Solar Energy Assistance

Neeraj Budhraja (neeraj_budhraja@yahoo.com)
DTU

Research Article

Keywords: Biodiesel, Linseed Oil, Taguchi's Method, Solar Paraboloid Dish, Response Surface Methodology

Posted Date: August 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1977661/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Biodiesel is a biofuel produced from vegetable oils and animal fats. The study describes the solar-assisted biodiesel production from linseed oil and the parameter optimization using Taguchi's L_{27} orthogonal approach and response surface methodology (RSM). A solar paraboloid dish of collector area 6.1 m^2 and concentration ratio approx. 200 is used for the transesterification process. The yearly and daily solar radiation data shows that May-June has longer solar radiation availability during the daytime; and are favorable months for experimentation. The results show that Taguchi's approach gives a maximum biodiesel yield of 89.14%, while the RSM model offers a slightly higher 91.9% yield. However, the RSM analysis predicted 91.1% (maximum biodiesel yield) at molar ratio (MR) 8.92:1, reaction time (RT) 108.97 minutes, and catalyst concentration (CC) 0.61 wt.%, respectively. The ANOVA analysis found that the MR has the highest percentage contribution of 75.67%, followed by CC (15.9%) and RT (5.69%). The biodiesel composition is determined using gas chromatography, and the various other fuel properties are measured as per ASTM testing methods. The study successfully confirms the solar heating usage for the transesterification process.

1. Introduction

Energy is essential for developing infrastructure, industries, transportation, and meeting other basic needs. The energy is mainly extracted from fossil fuels like crude oil and coal. The extensive crude oil usage has become a significant economic problem for countries with lower crude oil reserves [1–3]. This situation has led to higher crude oil prices; thus, making a disbalance in the world's economy, but it has no impact on demand which is breaking records year by year. The crude oil burning also adds toxic gases to the air leading to increased annual death rates due to air pollution. These reasons are behind restricting and minimizing crude oil usage [4–6]. Therefore, the other compatible crude oil alternatives are being explored for many years; and biodiesel is one such alternative [6–8].

Biodiesel is an alkyl ester produced by the reaction between vegetable oil and alcohol in the presence of a catalyst [9]. Biodiesel has a very vast history; however, the higher production cost, raw material availability, and lower cost of available conventional crude oil never let it become a potential source to replace diesel [6]. Many researchers have tried different biodiesel extraction techniques to cope with these difficulties and worked with various edible and non-edible oils [10–13]. They also explored other sources for biodiesel production, including waste cooking oil, animal fats, and microalgae [14–16]. The higher free fatty acid (FFA) content in oils reduces the biodiesel yield, and thus, it needs to be lowered before the biodiesel production process. Therefore, a two-stage transesterification is performed to reduce the FFA content in the oil [8, 17].

Various techniques are available for biodiesel production. Mechanical stirring is the most familiar technique for biodiesel production. It is a less efficient technique because it gives comparably lower biodiesel yield than other techniques and consumes much energy, while the reaction time is in hours [2, 15, 18]. Another technique is an ultrasonic-assisted method. It uses ultrasonic waves to generate a

cavitation effect within the mixture, which initiates the transesterification reaction faster and reduces the reaction time to 15–20 minutes with an improved biodiesel yield to about 92–95%. However, it demands tremendous energy to generate ultrasonic waves [15, 19]. The microwave-assisted method is another technique used to produce biodiesel. It is an advanced technology that gives a biodiesel yield of up to 96–99% within 5–7 minutes [20–22]. However, the higher energy demand and smaller batch production make it impracticable for commercial-scale production. Researchers and scientists are working on different techniques capable of continuous biodiesel production from various feedstocks at a lower cost [23–25].

Maintaining the temperature for transesterification reaction consumes the maximum energy in the biodiesel production process. Thus, renewable energy sources such as solar heat or waste heat from power plants must be considered [27–29]. Solar energy is dynamic in nature and available only during the daytime. Figure 1 shows the average monthly solar radiation availability and clearness index of Northern India throughout the year. The daily radiation from March to June is excellent for utilizing maximum solar energy, and different regions can be considered accordingly. Solar heat energy can efficiently replace conventional heating sources for biodiesel production [30, 31]; however, the solar devices should be chosen according to the temperature required for the reaction [32, 33]. The waste heat from the power plant is also a good heating source, but transferring it to a distant location is a difficult task [34].

In this experimental study, the solar-assisted method produces biodiesel from *Linum usitatissimum* (linseed oil), and the parameters are optimized using Taguchi's L_{27} orthogonal approach that are further optimized for enhanced biodiesel yield using response surface methodology (RSM). A solar parabolic dish of an area of 6.1 m^2 and approx. 200 concentration ratio is used to provide heating for the transesterification process. The study aims to find solar heat energy's potential to replace conventional heating sources in biodiesel production process. The current experimentation is for small batch-type biodiesel production; however, the study explored the possibility of using solar heat energy in a commercial-scale biodiesel production process. Therefore, the future work may focus on designing a continuous biodiesel production equipment compatible with solar energy.

2. Materials And Method

2.1 Solar Survey

Data for solar radiation with cloud clearness is collected every month throughout the year from earlier published works [26, 35]. It helps determine the suitable months for experimental work in Delhi Technological University, Delhi, India (28.7497° North and 77.1188° East). The study shows that the second half of May (15th -31st May) to the first half of June (1st-15th June) is the best possible period to perform experimentations because the maximum sunshine days are available during this period.

Solar radiation behavior is understood by a Tenmars TM-207 solar power meter (Fig. 2), which measures solar irradiation falling on earth. The reading is collected at every 30 minutes throughout the day before and between the experimentation. The solar radiation graph shows a similar trend for each day, rising from morning to afternoon and then declining (as shown in Fig. 3). Thus, the day is divided into three sessions, namely, morning (10:30 am to 12:30 pm), afternoon (12:30 pm to 2:30 pm), and post-afternoon (2:30 pm to 4:30 pm); and every experiment is performed in all three sessions of the day randomly, and the average reading is considered for calculations.

2.2 Materials

The matured seeds of *Linum usitatissimum* (linseed) are collected from the local market nearby the college campus of Delhi Technological University, Delhi, India, and the oil extraction is done by mechanical pressing [36]. Thus, the linseed oil is chosen as the raw material for biodiesel (methyl ester) production. Alcohol used in the process is industrial-grade methyl alcohol, while the catalyst taken is potassium hydroxide (KOH) with 56.11 molecular weight. The ranges of process parameters are chosen from the earlier published works [37–39], i.e., alcohol/oil ratio or molar ratio (MR) 6:1, 7.5:1 and 9:1, and catalyst concentration (CC) 0.5, 0.75 and 1.0 wt.%. However, the reaction time (RT) chosen are 90, 105 and 120 minutes taking into account the dynamic nature and availability of solar radiation throughout the day.

2.3 Biodiesel production method

Raw linseed oil is initially heated in a vessel at about 110°C using a conventional heater for approximately 10 minutes for moisture removal because moisture may lead to the saponification process, which eventually reduces the biodiesel yield [13]. The warm raw linseed oil is then allowed to cool down below 65°C before adding the catalyst-alcohol solution. Meanwhile, a methanol and catalyst KOH solution is prepared by mechanically dissolving KOH in methanol. The catalyst-alcohol solution is then poured into the raw linseed oil, and the container is sealed to prevent methanol vapors escape during the transesterification process. An aluminium foil covers the sealed container to restrict the UV radiation entering the container, which may cause radical formation. The aluminium foil also provides uniform heating to the sample. Next, the container is placed at the container stand (the focus of the solar receiver) to allow solar radiation to strike the aluminium foil and provide heat for the transesterification process. The solar receiver used is a solar paraboloid dish shown in Fig. 4, having a 6.1 m^2 collector area with a maximum temperature generation of 115ºC at 900–1000 W/m^2 solar irradiance and a concentration ratio of approx. 200. During the experiments, about 600–850 W/m^2 solar irradiance generating a temperature range between 75°C and 95°C is attained inside the container.

Once a sample is processed, it is transferred to a conical separating funnel and then kept undisturbed for about 8–10 hours to settle heavy glycerol in the sample. After the glycerol settles, it is taken out from the sample. The remaining liquid still contains dissolved KOH and traces of glycerol, which is removed by the water washing process. The water washing is performed by adding warm water to the biodiesel and

allowing it to settle for another 3–4 hours. The KOH dissolved water is then taken out, and the rest of the biodiesel is again heated to 105℃ using a conventional heater for removing any moisture left out during water washing [40]. The final product obtained is pure linseed methyl ester (LSME) or linseed biodiesel.

2.4 Design of Experiment

The robust designs for statistical calculations are developed using Taguchi's methods for engineering problems. These methods are simple and efficiently optimize the process parameters. An orthogonal L_{27} array is framed to assess the effects of various biodiesel production parameters at their different levels. The chosen parameters and their levels are represented in Table 1. It presents the complete evidence of parameters influencing the biodiesel yield [41, 42].

Table 1
Parameters and their levels for orthogonal L_{27} Taguchi's approach

Design Factors		Units	Level		
			1	2	3
A	Molar Ratio (MR)	—	6:1	7.5:1	9:1
B	Reaction Time (RT)	min.	90	105	120
C	Catalyst Concentration (CC)	wt.%	0.50	0.75	1.00

The signal to noise (S/N) ratio measures the experimental deviation to the desired value of performance parameters. The three types of S/N ratio are Larger the Better; Smaller the Better; and Nominal the Better [42, 43]. For biodiesel production, the Larger the Better S/N ratio is used. Here, the result is computed to get the maximum value of the response (biodiesel yield). The equation used in the process is as given by Eq. 1: -

$$S/Nratio = -10log_{10}\left\{\frac{1}{n}\sum_{j=1}^{N}\left(\frac{1}{Y_j^2}\right)\right\}$$

1

where n = number of responses in the factor level combination; and

Y_j = responses for the given factor level combination

3. Results And Discussion

Biodiesel Production

Current work focuses on maximizing biodiesel yield using Taguchi's approach and response surface methodology (RSM). In Table 2, the rank of all the parameters is computed along with the S/N ratios at different levels. Here, the delta value is used for determining the ranks of performance parameters, the highest delta value denotes rank 1, and so on. Hence, the molar ratio with rank 1 is identified as the most effective performance parameter in biodiesel production.

Table 2. Response table of operating parameters

Parameters	Level			Delta	Rank
	1	2	3		
Molar Ratio (oil/methanol)	38.17	38.43	38.79	0.60	1
Reaction Time (minutes)	38.38	38.55	38.48	0.23	3
Catalyst Concentration (wt.%)	38.59	38.48	38.31	0.13	2

Table 3 depicts the biodiesel yield (in percentage) obtained for each run (or experiment) in each day session. Data shows less than 1% deviation in biodiesel yield; thus, the mean value of three sessions is considered for further calculations. The mean biodiesel yield (in percentage) at various experiments (or run) and the S/N ratios are also given in Table 3. The mean biodiesel yield ranges from 77.45% (minimum) to 89.14% (maximum), and similar results are obtained by *Kumar et al.* (2017) [42] and *Kumar et al.* (2018) [43]. The highest mean yield of 89.14% is obtained in run 26 by MR 9:1, RT 120 mins. and CC 0.75 wt.%. Also, the optimum condition, which is measured by the highest value of the S/N ratio (38.94), is achieved by two experiments, i.e., run 22 and run 26, respectively.

Table 3. Biodiesel yield and S/N ratio of L_{27} experiments performed

Run	Molar Ratio	Reaction Time (min.)	Catalyst Conc. (wt. %)	Biodiesel Yield (%)			Mean Yield (%)	S/N ratio
				Morning	Afternoon	Post-Afternoon		
1	6:1	90	0.5	80.23	80.32	80.29	80.28	38.18
2	6:1	90	0.75	81.22	81.29	81.21	81.24	38.15
3	6:1	90	1	77.43	77.44	77.48	77.45	37.74
4	6:1	105	0.5	84.59	84.67	84.57	84.61	38.56
5	6:1	105	0.75	82.11	82.23	82.17	82.17	38.24
6	6:1	105	1	78.73	78.97	78.91	78.87	37.97
7	6:1	120	0.5	84.76	85.18	85.06	85.00	38.49
8	6:1	120	0.75	81.67	81.71	81.72	81.70	38.34
9	6:1	120	1	77.81	78.00	77.95	77.92	37.84
10	7.5:1	90	0.5	83.97	84.11	84.13	84.07	38.42
11	7.5:1	90	0.75	81.59	81.76	81.78	81.71	38.30
12	7.5:1	90	1	83.08	83.11	83.17	83.12	38.42
13	7.5:1	105	0.5	86.34	86.43	86.49	86.42	38.75
14	7.5:1	105	0.75	82.57	82.67	82.71	82.65	38.33
15	7.5:1	105	1	84.89	84.91	84.96	84.92	38.58
16	7.5:1	120	0.5	83.58	83.57	83.62	83.59	38.50
17	7.5:1	120	0.75	82.12	82.16	82.23	82.17	38.26
18	7.5:1	120	1	82.19	82.27	82.32	82.26	38.28
19	9:1	90	0.5	85.99	86.12	86.19	86.10	38.69
20	9:1	90	0.75	88.31	88.34	88.34	88.33	38.92
21	9:1	90	1	85.29	85.33	85.40	85.34	38.64
22	9:1	105	0.5	88.73	88.79	88.79	88.77	38.94
23	9:1	105	0.75	87.14	87.12	87.16	87.14	38.87
24	9:1	105	1	86.83	86.78	86.76	86.79	38.73
25	9:1	120	0.5	86.94	87.00	86.94	86.96	38.83
26	9:1	120	0.75	89.11	89.16	89.15	89.14	38.94
27	9:1	120	1	84.51	84.55	84.50	84.52	38.56

Initially, the effects of operating parameters on the response are analyzed using Taguchi's approach, and Fig. 5 shows the impact of different operating parameters on the biodiesel yield at all three levels. The peak value defines the optimum value of all parameters for different S/N ratio plots. Consequently, MR 9:1 (level 3), RT 105 min. (level 2), and CC 0.5 wt.% (level 1) achieves the highest biodiesel yield. It

indicates that the above process parameters are required to obtain the maximum biodiesel yield at the optimized conditions.

Regression analysis is performed by the data acquired in biodiesel yield from eq. 2, and the percentage of the coefficient of regression (97.72%) is given by eq. 3: -

$$\text{Yield} = 5.9 + 1.72\ MR + 1.379\ RT + 1.79\ CC + 0.584\ (MR)^2 - 0.00588\ (RT)^2 - 0.306\ (CC)^2 - 0.0267\ MR*RT + 0.873\ MR*CC - 0.0351\ RT*CC \quad (2)$$

$$R^2 = 97.72\% \quad (3)$$

Experimental and predicted values of biodiesel yield and the error percentage in different experiments of Taguchi's L_{27} approach are given in Table 4. The result shows a very close relationship between the experimental and predicted values of biodiesel yield. The experimental yield values (%age) show a minimal deviation from predicted yield values (%age), and the difference is measured as the mean error percentage for each experimentation. The highest mean error percentage is about 1.17%, which indicates that the experimental yield values are not more than 1.2% deviated from the predicted yield values.

Table 4. Experimental and predicted values for Taguchi's L_{27} approach

Experiment (Run)	Molar Ratio	Reaction Time (min.)	Catalyst Conc. (wt. %)	Yield (%)		Error (%)
				Experimental	Predicted	
1	6:1	90	0.5	80.28	81.07	0.9884
2	6:1	90	0.75	81.24	80.81	-0.5186
3	6:1	90	1	77.45	77.07	-0.4807
4	6:1	105	0.5	84.61	84.79	0.2045
5	6:1	105	0.75	82.17	81.64	-0.6481
6	6:1	105	1	78.87	79.23	0.4558
7	6:1	120	0.5	85.00	84.04	-1.1371
8	6:1	120	0.75	81.70	82.65	1.1675
9	6:1	120	1	77.92	77.93	0.0163
10	7.5:1	90	0.5	84.07	83.37	-0.8268
11	7.5:1	90	0.75	81.71	82.19	0.5908
12	7.5:1	90	1	83.12	83.33	0.2556
13	7.5:1	105	0.5	86.42	86.55	0.1503
14	7.5:1	105	0.75	82.65	82.48	-0.1973
15	7.5:1	105	1	84.92	84.95	0.0390
16	7.5:1	120	0.5	83.59	84.16	0.6761
17	7.5:1	120	0.75	82.17	81.85	-0.3890
18	7.5:1	120	1	82.26	82.01	-0.2984
19	9:1	90	0.5	86.10	86.00	-0.1143
20	9:1	90	0.75	88.33	88.27	-0.0696
21	9:1	90	1	85.34	85.50	0.1873
22	9:1	105	0.5	88.77	88.47	-0.3413
23	9:1	105	0.75	87.14	87.84	0.7983
24	9:1	105	1	86.79	86.40	-0.4524
25	9:1	120	0.5	86.96	87.36	0.4616
26	9:1	120	0.75	89.14	88.50	-0.7114
27	9:1	120	1	84.52	84.75	0.2754

Analysis of variance (ANOVA)

ANOVA determines the significance of various parameters in the process. The ANOVA analysis (as represented in Table 5) defines the importance of process parameters for biodiesel yield. The significance of a model is determined by its p-value (<0.0500), where a higher p-value states the process parameter is insignificant. In this case, MR and CC have a p-value of 0.000, making them highly significant model

terms, while RT has a p-value of 0.011; thus, it is also a significant model term but less effective than MR and CC.

Table 5. ANOVA table for biodiesel yield

Source	Degree of Freedom	Seq. SS	Adj. SS	Adj. MS	F-value	P-value	Contribution (%)
MR	2	163.115	163.115	81.5573	110.46	0.000	75.67
RT	2	12.262	12.262	6.1308	8.30	0.011	5.69
CC	2	34.278	34.278	17.1391	23.21	0.000	15.90
MR*RT	4	4.834	4.834	1.2084	1.64	0.256	---
MR*CC	4	29.056	29.056	7.2641	9.84	0.004	---
RT*CC	4	9.320	9.320	2.3300	3.16	0.078	---
Residual Error	8	5.907	5.907	0.7383			2.74
Total	26	258.771					100

Percentage contribution of MR, RT, and CC is also computed using the ANOVA table (Table 5), where Seq. SS terms are considered for calculating the percentage contribution of a parameter in the process. The highest contributing parameter is the molar ratio (75.67%), followed by catalyst concentration (15.9%) and reaction time (5.69%), while a residual error of 2.74% is also observed. It indicates that a slight variation in the molar ratio will significantly affect biodiesel yield.

Response surface methodology (RSM)

RSM is based upon the different mathematical and statistical practices established through the experimental design of adequate empirical models. It shows the relationship between the observed and theoretically obtained values from different empirical models. Fig. 6 depicts a connection between predicted values and actual values for biodiesel yield in percentage; it also shows the closeness of the two values.

Response surface methodology (RSM) also gives the interactive effects between different factors in surface-contour plots, which helps better understand the effects of two factors on response [44]. The surface-contour plots for the above experiments are shown in Fig. 7, 8, and 9. In Fig. 7, the effects of MR and CC are depicted, while the RT is taken as 105 min. (actual factor). The value of RT is taken from Taguchi's approach, which gives optimized biodiesel yield at 105 min. The figure shows that biodiesel production enhances when MR (alcohol to oil) is higher that is similar to the results obtained by *Fan et al. (2011)* [44]. Consequently, the catalyst concentration between 0.6 and 0.9 wt.% shows a slight decline,

going against the findings of *Yadav et al.* (2018) [45]. However, the plot shows the potential for higher biodiesel yield with the rise in molar ratio beyond 9:1.

The interactive effects between CC and RT are represented in Fig. 8. Here, the curve's flatness shows a lower percentage contribution of these parameters, as described in the ANOVA table (table 5). CC and RT's percentage contributions are less than MR; thus, putting molar ratio to 9:1 (actual factor) will show less fluctuation in the surface counter curve. The transesterification process in solar-assisted biodiesel production takes more time than conventional methods; as a result, the optimized RT is observed between 95 and 110 mins. However, after 110 mins, a decline in biodiesel yield is observed due to reverse reaction. Similar observations are also noticed by *Mihankhah et al.* (2016) [12].

In Fig. 9, the interaction of RT with MR is shown for biodiesel yield (as a response). As discussed, the biodiesel yield rises with the molar ratio; thus, a similar trend is observed in the figure. Higher molar ratio with RT more than 95 min. gives more than 83% of biodiesel yield. From different published works, it was observed that the transesterification process requires a 3:1 M ratio (alcohol to triglyceride) for alkyl ester (biodiesel) conversion [46, 47]; however, to get a biodiesel yield greater than 80%, an MR of 9:1 or more is favorable. In contrast, the higher MR results in extra alcohol content, creating difficulty in glycerol separation. Thus, a relevant MR between 9 and 12 is acceptable for biodiesel production, while a higher MR of 15:1 reduces the biodiesel yield due to difficulty in glycerol recovery [33, 45]. RSM analysis concluded that the optimized parameters for desirable biodiesel yield (91.1%) are molar ratio 8.92:1, reaction time 108.97 minutes, and catalyst concentration 0.61 wt.%, respectively. Five consecutive experiments are performed to confirm the values obtained through the RSM model, and the average of the five experiments is found to be 91.9% biodiesel yield.

Fuel properties

For considering any fuel for internal combustion (IC) engines, the fuel properties must match the engine input parameters as per ASTM standards. Thus, the desirable fuel properties are measured for biodiesel (LSME) and raw linseed oil, and compared with diesel. Table 6 compares desirable fuel properties of LSME and diesel as per ASTM testing methods.

Table 6. Desirable fuel properties for CI engine

Properties	ASTM Testing method	SI Unit	Diesel	Raw linseed oil	Linseed methyl ester (LSME)		
					This work	[48]	[49]
Density at 15°C	D-1298	kg/m^3	809.6	914.0	869.8	860	872
Kinematic viscosity at 40°C	D-445	$\times 10^{-6}$ m^2/s	2.815	27.317	5.453	5	8.2
Calorific value	D-240	$\times 10^3$ kJ/kg	45.5	35.2	36.7	35.6	37.5
Flash point	D-93	°C	72	>210	162	187	161
Pour point	D-97	°C	-9	-15	-11	---	---
Cetane index	D-4737	---	48.1	40.2	42.7	---	55

The density of the conventional diesel is 809.6 kg/m^3, whereas raw linseed oil has a higher density of 914.0 kg/m^3. After the transesterification of linseed oil, the linseed methyl ester (LSME) is produced [50], and the density is reduced to 869.8 kg/m^3, which is under the ASTM standards [1]. In contrast, the kinematic viscosity of LSME is found to be 5.453×10^{-6} m^2/s which is twice of diesel (2.815×10^{-6} m^2/s) and is close to the upper limit of the ASTM standards [1, 13].

The calorific value (CV) of a fuel is the heating capacity of the fuel or the energy released when the fuel is burnt. The CV of diesel is measured maximum (45.5×10^3 kJ/kg), and LSME is measured about 36.7×10^3 kJ/kg. Flash point and pour point of LSME are 162°C and -11°C, while the cetane index is 42.7, respectively. Hence, it is concluded that the desirable properties of biodiesel are comparable to diesel (under ASTM standards) [1, 13]. As a result, a biodiesel blend with diesel could be used in the existing CI engines with no modifications, or pure biodiesel could directly be used with few modifications in the CI engine.

Fatty acid methyl ester (FAME) composition

The FAME composition of the oil is an essential parameter. It gives the composition of all the FAME content present in the biodiesel. Ultima make series 2100 gas chromatography (specifications in Table 7) is used to measure the FAME composition of biodiesel (LSME).

Table 7. Specifications of gas chromatography

Make	Ultima
Model	Series 2100
Detector	Flame ionization detector (FID)
Carrier gas	Nitrogen
Other gases	Hydrogen and zero air
Oven temperature	240°C
Rate of temperature rise	10°C/min

The gas chromatogram for LSME is shown in Fig. 10. The curve shows the different FAME compositions in percentage, while Table 8 presents the FAME composition of linseed methyl ester (LSME).

Table 8. FAME composition in wt.%

Fatty acid	Chemical formula	Degree of Saturation	Linseed methyl ester		
			This work*	[51]*	[52]*
Palmitic	$C_{16}H_{32}O_2$	16:0	5.41	6.58	5.69
Stearic	$C_{18}H_{36}O_2$	18:0	4.94	4.43	5.58
Oleic	$C_{18}H_{34}O_2$	18:1	18.73	18.51	20.59
Linoleic	$C_{18}H_{32}O_2$	18:2	15.95	17.25	15.80
Linolenic	$C_{18}H_{30}O_2$	18:3	54.12	53.21	51.38
Total saturated	-	-	10.45	11.01	11.90
Total unsaturated	-	-	89.35	88.99	88.10

*All values are in wt.%

The linolenic acid content is highest (54.12%) in LSME, whereas other higher fatty acids are linoleic acid (15.95%) and oleic acid (18.73%). However, a small quantity of palmitic acid (5.41%) and stearic acid (4.94%) is also present in the linseed methyl ester (as depicted in Table 8). The experimentally obtained FAME values are comparable to earlier work, and it is found that the values obtained (Table 8) do not differ significantly from other published data [36, 51–53].

Comparative analysis of results

A detailed comparison of various biodiesel production methods with feed oil is represented in Table 9. It is concluded that solar-assisted biodiesel production is well competitive with other techniques in biodiesel yield. However, the zero-power requirement in the solar-assisted process makes it superior to other methods.

Table 9. Comparative analysis of present work and other published works

Reference	Feed oil	Methods used	Power	Biodiesel yield
Predicted	Linseed oil	---	---	91.1%
Experimental (RSM model)	Linseed oil	Solar-assisted	0	91.9%
Experimental (ANOVA)	Linseed oil	Solar-assisted	0	91.6%
Mohite et al. [13]	Karanja-Linseed oil mixture	Mechanical starring and conventional heating	180 W	78.9%
Singh et al. [15]	Microalgae oil	Ultrasonic	50 W	98%
Uddin et al. [16]	Waste cooking oil	Mechanical starring and conventional heating	210 W	79%
Hsiao et al. [20]	Soybean oil	Microwave-assisted	300 W	96.6%
Kumar et al. [21]	Pongamia pinnata seed oil	Microwave-assisted	300 W	96%
Mohan et al. [40]	Semal oil	Ultrasonic cavitation and mechanical stirring	780 W	90.7%
Kumar et al. [43]	Pongamia oil	Mechanical starring and conventional heating	180 W	90.2%

*Bold is used to represent present work results

4. Conclusions

Biodiesel is produced from the reaction between Linum usitatissimum oil (linseed oil) and methanol in the presence of KOH. A solar paraboloid dish of a 6.1 m^2 collector area and approx. 200 concentration ratio provides heat energy for the transesterification process. An L_{27} orthogonal array is designed using Taguchi's approach with the molar ratio, reaction time, and catalyst concentration at three different levels. RSM and ANOVA analysis are performed to optimize process parameters and determine their percentage contribution in biodiesel production. The confirmation experiment is performed to validate the optimized parameter values obtained using RSM. The following findings are obtained from the study: -

- During the daytime, the solar paraboloid dish attained a temperature of 75–95°C at about 600–850 W/m^2 solar radiation.
- Taguchi's L_{27} orthogonal array experimentally achieves the highest mean biodiesel yield of 89.14%.
- FAME content in LSME is characterized using GC, and the various biodiesel properties are determined as per ASTM testing methods.

- Optimized parameters using RSM analysis are MR 8.92:1, RT 108.97 minutes, and CC 0.61 wt.% producing 91.1% of predicted biodiesel yield.
- Experimental biodiesel yield of 91.9% is achieved using the RSM model at optimized parameters and comparable with other published works.
- ANOVA analysis showed the maximum contribution of MR (75.67%), followed by CC (15.9%) and RT (5.69%), respectively.
- Taguchi's approach provides interaction of all parameters; however, the RSM analysis is better due to the higher interaction of all process parameters.

Thus, the study showed the excellent potential of using solar heat energy for biodiesel production, reducing the dependence on conventional heating sources.

5. Future Scope

The current study is performed in a small-scale batch-type biodiesel production process to understand the potential of using solar heat energy in the transesterification process. The study shows comparatively good biodiesel yield without using any conventional heating source during the transesterification process. However, commercial-scale biodiesel production requires an efficient solar capturing device. Thus, future research can focus on finding the potential of other solar heat capturing devices and raising the biodiesel yield for continuous biodiesel production at a commercial scale.

Declarations

Acknowledgment

This work is supported by the Delhi Technological University, Delhi, India. The authors are thankful for all the support provided by the IC Engine lab, Mechanical Engineering Department, DTU for conducting experiments.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflicts of interest/Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Consent for publication

Authors have agreed to submit it in its current form for consideration for publication in the journal.

Author Contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by *Neeraj Budhraja*. The first draft of the manuscript was written by *Neeraj Budhraja* and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Availability of data and material

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. Additional data (figures showing experimental setup and gas chromatography) is presented in the Supplementary file.

Ethics approval and consent to participate

Not applicable. No tests, measurements or experiments were performed on humans as part of this work.

References

1. Jakeria MR, Fazal MA, Haseeb ASMA (2014) Influence of different factors on the stability of biodiesel: A review. Renew Sustain Energy Rev 30:154–163. https://doi.org/10.1016/j.rser.2013.09.024
2. Yadav AK, Khan ME, Pal A, Dubey AM (2016) A comparative study on ultrasonic cavitation and mechanical stirring method towards efficient production of biodiesel from non-edible oils and performance testing on a CI engine. Int J Environ Waste Manag 18:349. https://doi.org/10.1504/IJEWM.2016.081836
3. Budhraja N, Pal A, Jain M, Mishra RS (2021) Comparative Analysis of the Engine Emissions from CI Engine Using Diesel–Biodiesel–Ethanol Blends. In: Kumar A, Pal A, Kachhwaha SS, Jain PK (eds). Springer Singapore, Singapore, pp 363–370
4. Brunet R, Antipova E, Guillén-Gosálbez G, Jiménez L (2012) Reducing the Environmental Impact of Biodiesel Production from Vegetable Oil by Use of a Solar-Assisted Steam Generation System with Heat Storage. Ind Eng Chem Res 51:16660–16669. https://doi.org/10.1021/ie301391h
5. Pal A, Kachhwaha SS (2013) Waste cooking oil: A promising feedstock for biodiesel production through power ultrasound and hydrodynamic cavitation. J Sci Ind Res (India) 72:387–392
6. Sajid Z, Khan F, Zhang Y (2016) Process simulation and life cycle analysis of biodiesel production. Renew Energy 85:945–952. https://doi.org/10.1016/j.renene.2015.07.046
7. Vasudevan P, Sharma S, Kumar A (2005) Liquid fuel from biomass: An overview. J Sci Ind Res (India) 64:822–831
8. Kombe GG, Temu AK, Rajabu HM, et al (2013) Low Temperature Glycerolysis as a High FFA Pre-Treatment Method for Biodiesel Production. Adv Chem Eng Sci 03:248–254. https://doi.org/10.4236/aces.2013.34032

9. Kant Bhatia S, Kant Bhatia R, Jeon JM, et al (2021) An overview on advancements in biobased transesterification methods for biodiesel production: Oil resources, extraction, biocatalysts, and process intensification technologies. Fuel 285:119117. https://doi.org/10.1016/j.fuel.2020.119117
10. Buasri A, Worawanitchaphong P, Trongyong S, Loryuenyong V (2014) Utilization of Scallop Waste Shell for Biodiesel Production from Palm Oil – Optimization Using Taguchi Method. APCBEE Procedia 8:216–221. https://doi.org/10.1016/j.apcbee.2014.03.030
11. Demirbas A, Bafail A, Ahmad W, Sheikh M (2016) Biodiesel production from non-edible plant oils. Energy Explor Exploit 34:290–318. https://doi.org/10.1177/0144598716630166
12. Mihankhah T, Delnavaz M, Khaligh NG (2016) Eco-friendly biodiesel production from olive oil waste using solar energy. Energy Sources, Part A Recover Util Environ Eff 38:3668–3672. https://doi.org/10.1080/15567036.2016.1167792
13. Mohite S, Kumar S, Maji S, Pal A (2016) Production of Biodiesel from a mixture of Karanja and Linseed oils: Optimization of process parameters. Iran J energy Environ 7:12–17. https://doi.org/10.5829/idosi.ijee.2016.07.01.03
14. Huang G, Chen F, Wei D, et al (2010) Biodiesel production by microalgal biotechnology. Appl Energy 87:38–46. https://doi.org/10.1016/j.apenergy.2009.06.016
15. Singh A, Pal A, Maji S (2017) Biodiesel production from microalgae oil through conventional and ultrasonic methods. Energy Sources, Part A Recover Util Environ Eff 39:806–810. https://doi.org/10.1080/15567036.2016.1263260
16. Uddin MR, Ferdous K, Uddin MR, et al (2013) Synthesis of Biodiesel from Waste Cooking Oil. Chem Eng Sci 1:22–26. https://doi.org/10.12691/ces-1-2-2
17. Berchmans HJ, Hirata S (2008) Biodiesel production from crude Jatropha curcas L. seed oil with a high content of free fatty acids. Bioresour Technol 99:1716–1721. https://doi.org/10.1016/j.biortech.2007.03.051
18. Yathish. K.V, Omkaresh. B. R RS (2013) Biodiesel Production Process Optimization from Sugar Apple Seed Oil (Annona squamosa) and Its Characterization. Int J Eng Res Technol 2:31–36
19. Yadav AK, Khan ME, Pal A (2017) Biodiesel production from oleander (Thevetia Peruviana) oil and its performance testing on a diesel engine. Korean J Chem Eng 34:340–345. https://doi.org/10.1007/s11814-016-0270-8
20. Hsiao M, Lin C, Chang Y (2011) Microwave irradiation-assisted transesterification of soybean oil to biodiesel catalyzed by nanopowder calcium oxide. Fuel 90:1963–1967. https://doi.org/10.1016/j.fuel.2011.01.004
21. Kumar R, Kumar GR, Chandrashekar N (2011) Microwave assisted alkali-catalyzed transesterification of Pongamia pinnata seed oil for biodiesel production. Bioresour Technol 102:6617–6620. https://doi.org/10.1016/j.biortech.2011.03.024
22. Yaakob Z, Sukarman IS, Kamarudin SK, et al (2008) Production of Biodiesel from Jatropha Curcas by Microwave Irradiation. 235–239

23. Azeem MW, Hanif MA, Al-Sabahi JN, et al (2016) Production of biodiesel from low priced, renewable and abundant date seed oil. Renew Energy 86:124–132. https://doi.org/10.1016/j.renene.2015.08.006
24. Silitonga AS, Mahlia TMI, Kusumo F, et al (2019) Intensification of Reutealis trisperma biodiesel production using infrared radiation: Simulation, optimisation and validation. Renew Energy 133:520–527. https://doi.org/10.1016/j.renene.2018.10.023
25. Suresh T, Sivarajasekar N, Balasubramani K (2021) Enhanced ultrasonic assisted biodiesel production from meat industry waste (pig tallow) using green copper oxide nanocatalyst: Comparison of response surface and neural network modelling. Renew Energy 164:897–907. https://doi.org/10.1016/j.renene.2020.09.112
26. Usman M, Khan MT, Rana AS, Ali S (2018) Techno-economic analysis of hybrid solar-diesel-grid connected power generation system. J Electr Syst Inf Technol 5:653–662. https://doi.org/10.1016/j.jesit.2017.06.002
27. Aro E (2016) From first generation biofuels to advanced solar biofuels. Ambio 45:24–31. https://doi.org/10.1007/s13280-015-0730-0
28. Hallenbeck PC, Grogger M, Mraz M, Veverka D (2016) Solar biofuels production with microalgae. Appl Energy 179:136–145. https://doi.org/10.1016/j.apenergy.2016.06.024
29. Hou Z, Zheng D (2009) Solar utility and renewability evaluation for biodiesel production process. Appl Therm Eng 29:3169–3174. https://doi.org/10.1016/j.applthermaleng.2009.04.020
30. Bharti R, Guldhe A, Kumar D, Singh B (2020) Solar irradiation assisted synthesis of biodiesel from waste cooking oil using calcium oxide derived from chicken eggshell. Fuel 273:117778. https://doi.org/10.1016/j.fuel.2020.117778
31. Kumar R, Pal P (2021) Lipase immobilized graphene oxide biocatalyst assisted enzymatic transesterification of Pongamia pinnata (Karanja) oil and downstream enrichment of biodiesel by solar-driven direct contact membrane distillation followed by ultrafiltration. Fuel Process Technol 211:106577. https://doi.org/10.1016/j.fuproc.2020.106577
32. Bhattacharya S, Maurya R, Mishra SK, et al (2016) Solar driven mass cultivation and the extraction of lipids from Chlorella variabilis: A case study. Algal Res 14:137–142. https://doi.org/10.1016/j.algal.2016.01.009
33. Corro G, Tellez N, Bañuelos F, Eugenia M (2012) Biodiesel from Jatropha curcas oil using Zn for esterification step and solar radiation as energy source. Fuel 97:72–79. https://doi.org/10.1016/j.fuel.2012.02.029
34. León JA, Montero G, Coronado M, et al (2016) Solar Energy for a Solvent Recovery Stage in a Biodiesel Production Process. Int J Photoenergy 2016:1–7. https://doi.org/10.1155/2016/1048095
35. Masoom A, Kosmopoulos P, Bansal A, Kazadzis S (2020) Solar Energy Estimations in India Using Remote Sensing Technologies and Validation with Sun Photometers in Urban Areas. Remote Sens 12:254. https://doi.org/10.3390/rs12020254

36. Yadav AK, Dewangan A, Mallick A (2018) Synthesis and Stability Study of Biodiesel from Kachnar Seed Oil. J Energy Eng 144:04018053. https://doi.org/10.1061/(asce)ey.1943-7897.0000565
37. Ajala EO, Ajala MA, Ayinla IK, et al (2020) Nano-synthesis of solid acid catalysts from waste-iron-filling for biodiesel production using high free fatty acid waste cooking oil. Sci Rep 10:1–21. https://doi.org/10.1038/s41598-020-70025-x
38. Athar M, Zaidi S, Hassan SZ (2020) Intensification and optimization of biodiesel production using microwave-assisted acid-organo catalyzed transesterification process. Sci Rep 10:1–18. https://doi.org/10.1038/s41598-020-77798-1
39. Manojkumar N, Muthukumaran C, Sharmila G (2020) A comprehensive review on the application of response surface methodology for optimization of biodiesel production using different oil sources. J King Saud Univ - Eng Sci. https://doi.org/10.1016/j.jksues.2020.09.012
40. Mohan S, Pal A, Singh R (2017) The production of semal oil methyl esters through a combined process reactor. Energy Sources, Part A Recover Util Environ Eff 39:955–962. https://doi.org/10.1080/15567036.2016.1235059
41. Hariram V, Fernandes JL, Jaganathan R, et al (2017) Optimized Biodiesel Production and Emulsification of Pongamia Seed Oil Using Taguchi Method. Int J Renew Energy Res 7:1957–1961
42. Kumar N, Mohapatra SK, Ragit SS, et al (2017) Optimization of safflower oil transesterification using the Taguchi approach. Pet Sci 14:798–805. https://doi.org/10.1007/s12182-017-0183-0
43. Kumar GKBS, Rajesh K, Sharma AHK, et al (2018) Optimization of Biodiesel Production from Pongamia Oil using Taguchi Method. Int J Eng Res Technol 6:1–7
44. Fan X, Wang X, Chen F (2011) Biodiesel Production from Crude Cottonseed Oil: An Optimization Process Using Response Surface Methodology. Open Fuels Energy Sci J 4:1–8. https://doi.org/10.2174/1876973X01104010001
45. Yadav AK, Khan ME, Pal A, Singh B (2018) Ultrasonic-assisted optimization of biodiesel production from Karabi oil using heterogeneous catalyst. Biofuels 9:101–112. https://doi.org/10.1080/17597269.2016.1259522
46. Aslam MM, Khan AA, Cheema HMN, et al (2020) Novel mutant camelina and jatropha as valuable feedstocks for biodiesel production. Sci Rep 10:1–9. https://doi.org/10.1038/s41598-020-78680-w
47. Naveenkumar R, Baskar G (2020) Optimization and techno-economic analysis of biodiesel production from Calophyllum inophyllum oil using heterogeneous nanocatalyst. Bioresour Technol 315:. https://doi.org/10.1016/j.biortech.2020.123852
48. Kumar S, Jain S, Kumar H (2020) Experimental Study on Biodiesel Production Parameter Optimization of Jatropha-Algae Oil Mixtures and Performance and Emission Analysis of a Diesel Engine Coupled with a Generator Fueled with Diesel/Biodiesel Blends. ACS Omega 5:17033–17041. https://doi.org/10.1021/acsomega.9b04372
49. Akram W, Singh Y, Sharma A, Singh NK (2021) Experimental studies on performance and exhaust emission characteristics of a diesel engine fuelled with diesel-linseed oil methyl ester (LOME) blends.

Energy Sources, Part A Recover Util Environ Eff 43:754–768. https://doi.org/10.1080/15567036.2019.1632978

50. Dixit S, Kanakraj S, Rehman A (2012) Linseed oil as a potential resource for bio-diesel: A review. Renew Sustain Energy Rev 16:4415–4421. https://doi.org/10.1016/j.rser.2012.04.042
51. Popa V, Gruia A, Raba D-, et al (2012) Linseed Oil Composition - Popa. J Agroaliment Process Technol 18:136–140
52. Danish M, Nizami M (2019) Complete fatty acid analysis data of flaxseed oil using GC-FID method. Data Br 23:103845. https://doi.org/10.1016/j.dib.2019.103845
53. Lang X, Dalai AK, Bakhshi NN, et al (2001) Preparation and characterization of bio-diesels from various bio-oils. Bioresour Technol 80:53–62. https://doi.org/10.1016/S0960-8524(01)00051-7

Supplementary Material

Supplementary Information not available with this version.

Figures

Figure 1

Annual solar radiation and clearness index of Northern India [26]

Figure 2

Solar power meter

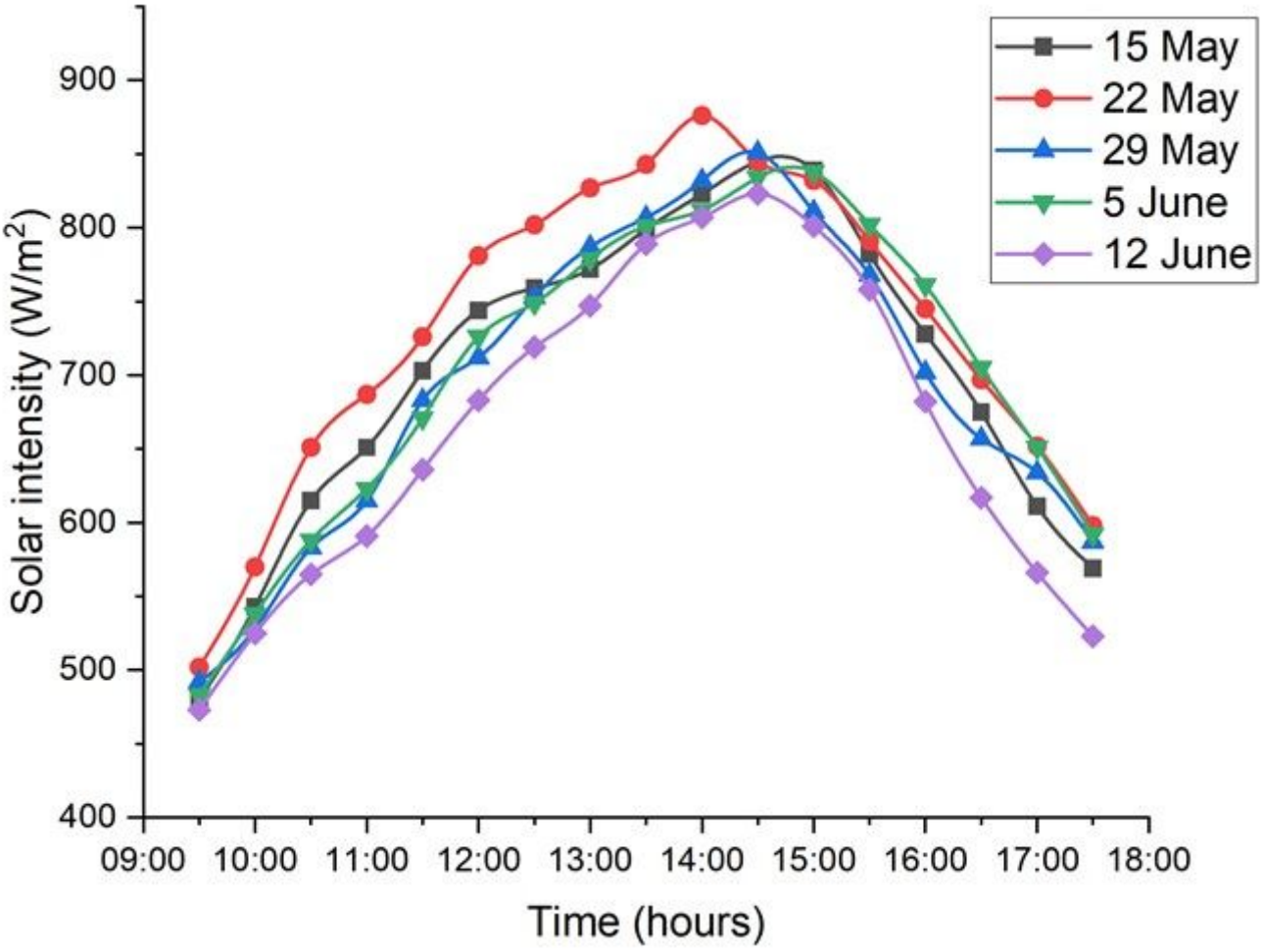


Figure 3

Daily solar radiation curve for Delhi

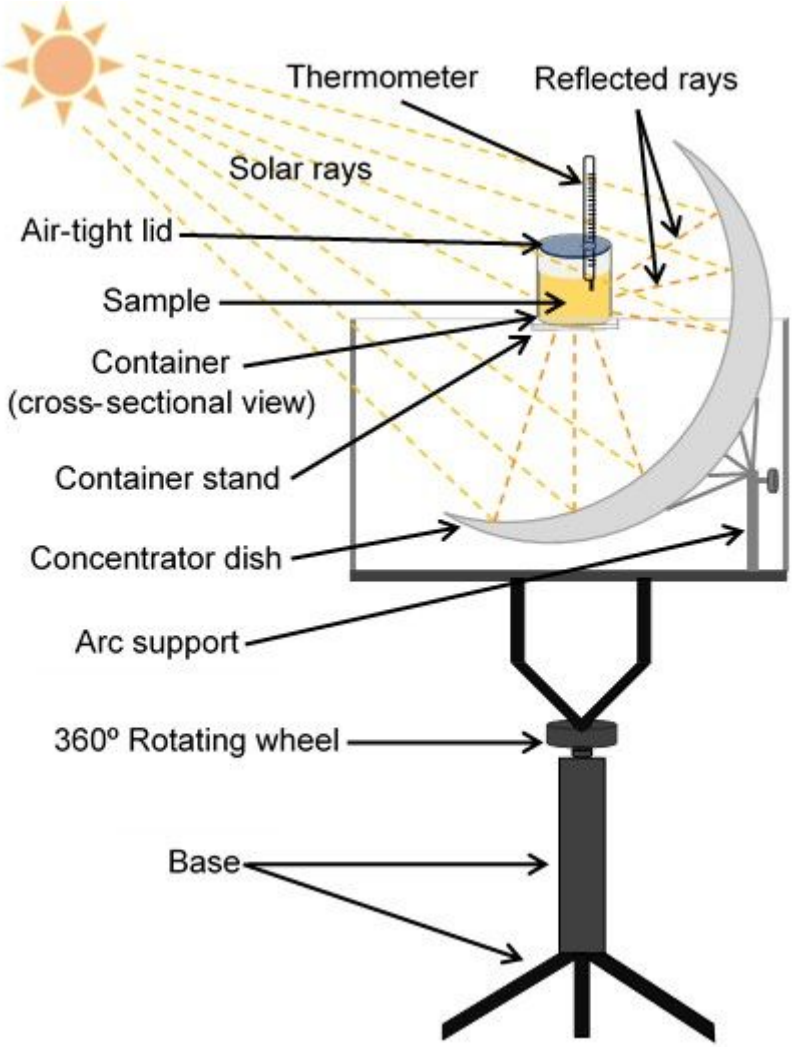


Figure 4

Solar paraboloid dish experimental setup

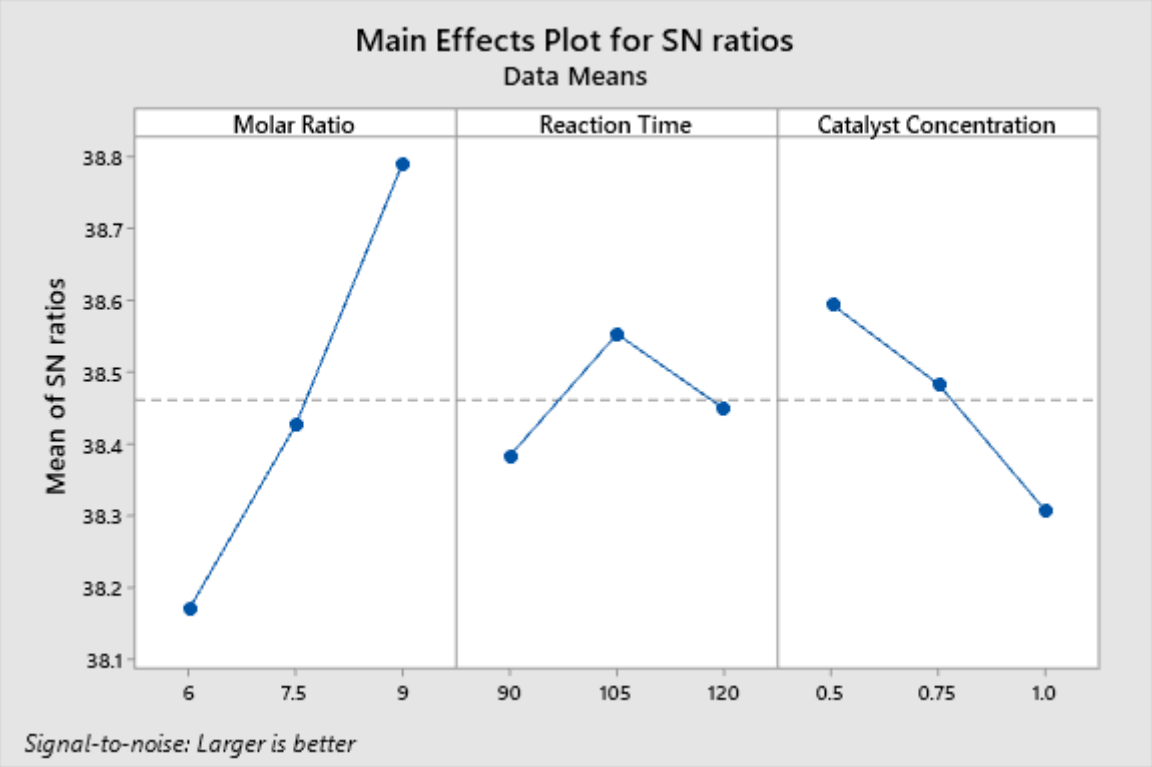


Figure 5

Effects of S/N ratios (Larger is Better) on various parameters

Figure 6

Predicted vs. actual value relationship for biodiesel yield percentage

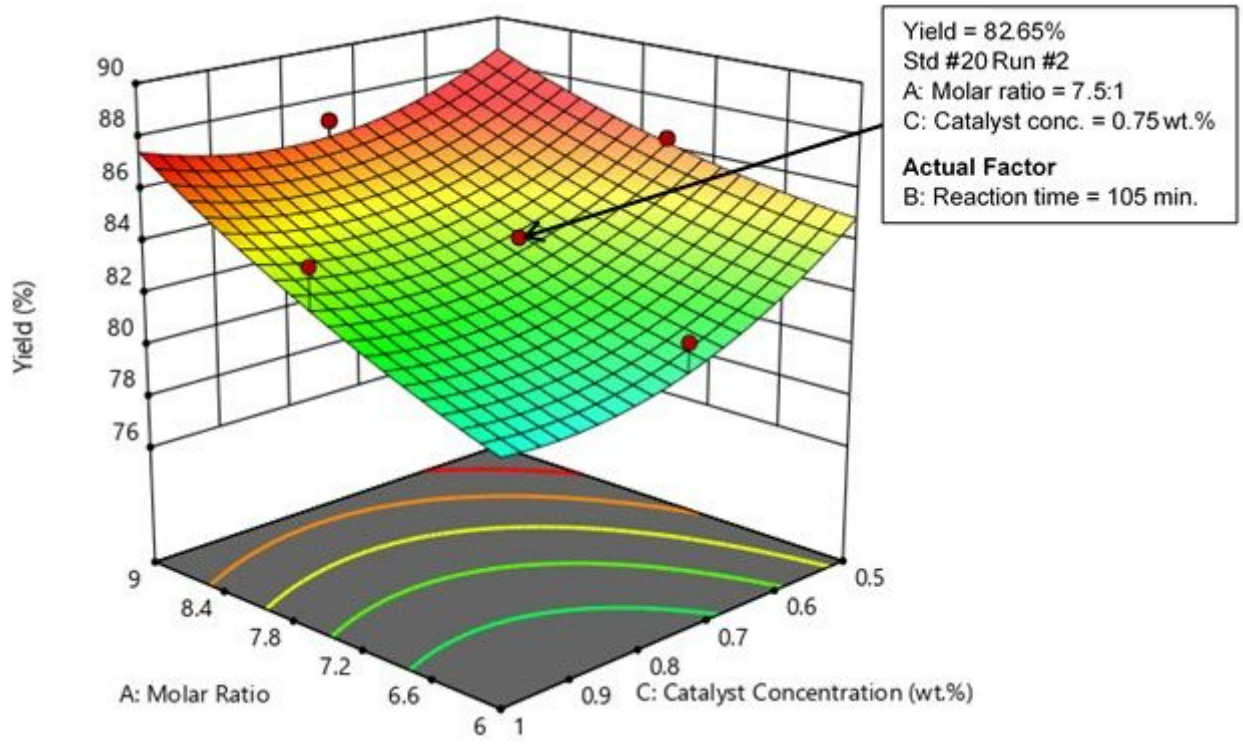


Figure 7

Interaction surface response curve between CC (wt.%) and MR

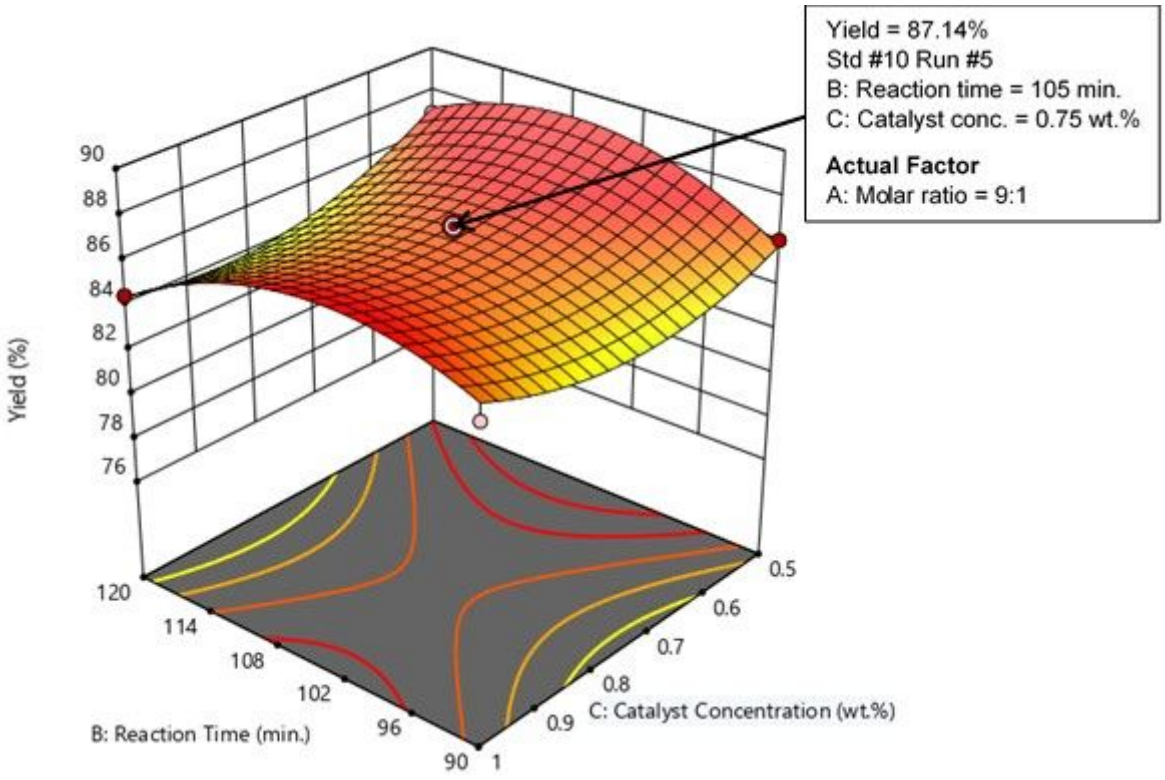


Figure 8

Interaction surface response curve between CC (wt.%) and RT (min.)

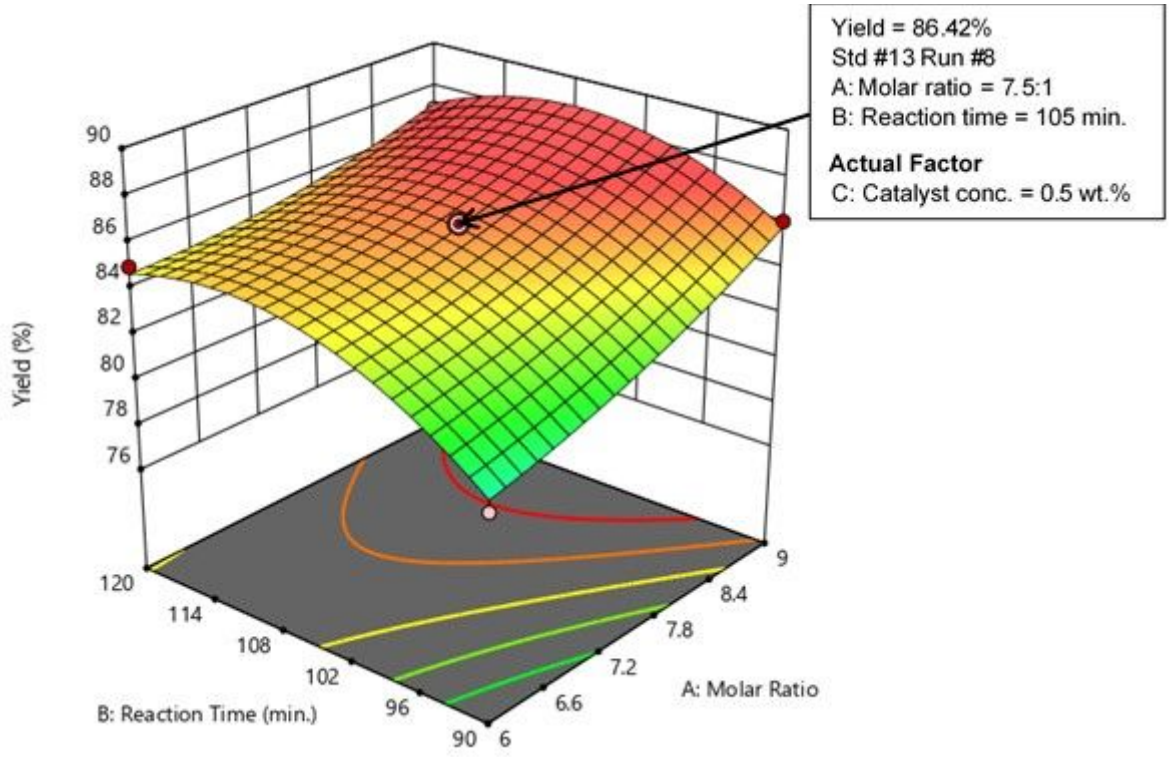


Figure 9

Interaction surface response curve between RT (min.) and MR (alcohol to oil)

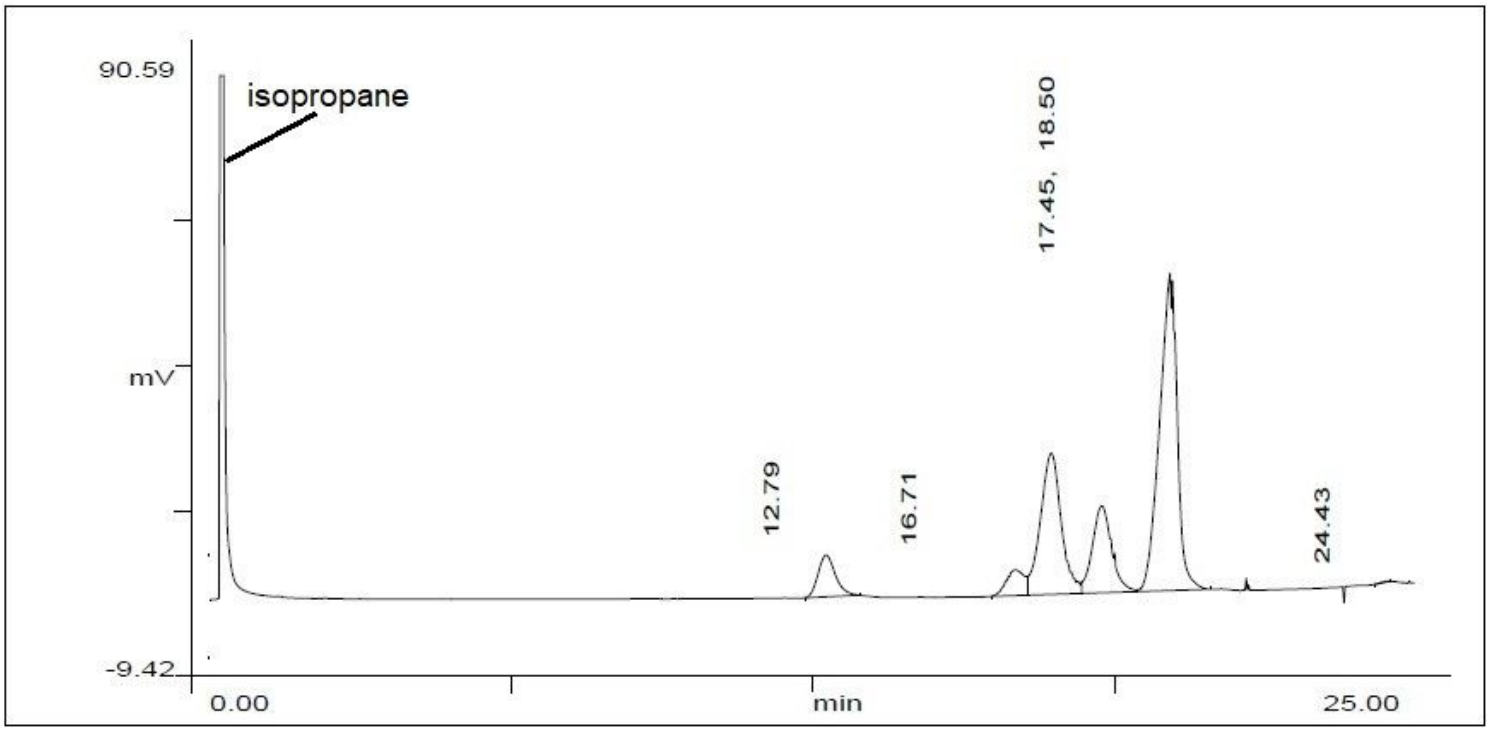


Figure 10

Gas chromatogram for linseed methyl ester (LSME)

Clean Technologies and Environmental Policy
https://doi.org/10.1007/s10098-022-02374-2

ORIGINAL PAPER

Performance evaluation of an industrial solar dryer in Indian scenario: a techno-economic and environmental analysis

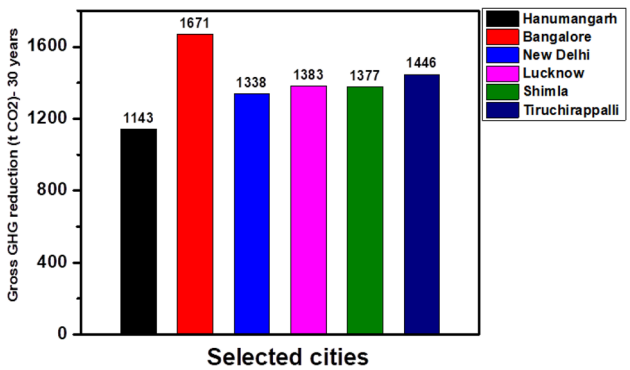
Ankit Srivastava[1,2] · Abhishek Anand[1] · Amritanshu Shukla[1,3] · Anil Kumar[4] · Atul Sharma[1]

Received: 14 February 2022 / Accepted: 20 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Food product drying is used to keep food products preserved for a long period and reduce post-harvest losses. Many industries dealing with food drying employ traditional drying systems that use traditional fuel for drying processes. Solar drying systems may be integrated with traditional industrial drying systems to save energy and reduce carbon emissions. The primary goal is to explore the technical, environmental, and economic feasibility of indirect type solar drying technology integrated with electric heaters as per the requirement of an industrial drying system in six cities throughout India. RETScreen solar drying model is intended to analyze energy production, solar fraction and savings of energy, costs, reductions in greenhouse gas emissions, financial feasibility, and various risk factors for the solar drying model in various locales. Sun drying project's economic viability is proven by measurements such as the simple payback period (SPP), internal rate of return (IRR), and equity payback period (EPP). This study's findings provide a clear vision for investors in the food drying sector to build their system in a specified area so that investors may maximize their profit while also supporting the environment by lowering carbon emissions. The simulation result of the presented scenario reveals that utilizing a 100 m^2 solar air heater annually saves the maximum energy in Bangalore (220,283 kWh), followed by Tiruchchirappalli (190,795 kWh) and Lucknow (182,486 kWh). Hanumangarh achieves a minimum energy savings of 150,548 kWh. Payback period for keeping a temperature of 40 °C with a 29.4 kg/s air flow rate is between 1 and 2 years. The largest value of EPP and SPP is found in Hanumangarh i.e., 0.56 years and 1.9 years respectively, while the minimum value of EPP and SPP is found in Bangalore i.e., 0.36 years and 1.3 years, respectively. The solar fraction can be enhanced further by expanding the surface area of the solar collector. It may also be concluded that the government must offer at least 50% incentives to make such installation practicable.

Graphical abstract



Keywords Industrial solar drying · Energy savings · GHG reduction · Financial analysis · RETScreen model

List of symbols

T_{PCF}	Positive cash flow year
C_t	After tax cash flow in t years
C_o	Initial investment required for the project
C_f	Cash flow per year
C_t	Net cash inflow during the period of time t
C	Initial cost

Extended author information available on the last page of the article

Published online: 02 August 2022

IG	Incentive and grants
C_E	Annual energy saving or income
C_C	Annual capacity saving or income
C_{RC}	Renewable energy production credit income
C_{GHG}	GHG reduction income
$C_{O\&M}$	Annual operations and maintenance costs
C_{FUEL}	Annual cost of fuel
IRR	Internal rate of return
NPV	Net present value
t	Time of project life

Introduction

Food and Agriculture Organization of United Nations (FAO) estimated that about one-third of the world's food is wasted or lost each year (Gustavsson et al. 2013). Since then, the global perception of the problem has changed fairly. By 2050, the world's population is predicted to reach 10.5 billion people, this loss of food, if prevented and effectively managed, can fulfill the food requirement of future generations. The largest post-harvest losses in developing countries are in perishables, such as fruits and vegetables. These losses probably reflect the poorly developed supply chain structure between agriculture and retail. For several crops, post-harvest losses in developing countries are relatively high. Major food loss occurs at the consumer level in developed nations, although some small part of food is wasted on-farm and between supply-chain phases. Field losses arise because farmers decide to turn down harvesting due to difficult market conditions. On the other hand, losses in developing countries emerge mostly at the field-to-market stage, with the consumer level responsible for the smallest losses. Insufficient storage facilities, early harvesting, a scarcity of infrastructure, insufficient retail facilities, and a scarcity of processing facilities are all factors that lead to high food losses in the food supply chain in developing countries (Aulakh et al. 2013).

Drying the crop is preferable for avoiding post-harvest crop loss and meeting future food needs. Food may be preserved for an extended period by drying the produce. Drying is a basic procedure in which the moisture content of a crop is reduced to the desired level, which varies depending on the crop by applying heat. Food that has been properly dried has a better nutritional value and may be fed to people. Because a lack of refrigerated storage facilities causes a substantial proportion of food to be lost in underdeveloped nations, drying can be employed to preserve food. Solar energy may be used to dry food since it is abundant. Open sun drying has certain drawbacks, such as the inability to manage temperature, which can cause food to get overheated at times, and the fact that it is a labor-intensive procedure involving many people. Furthermore, because of the great emphasis on the variability of weather conditions, drying in the open sun is quite difficult (Srivastava et al. 2021).

As the supply of conventional fuel decreases and the cost of fuel rises, solar energy may be a better choice to replace traditional fuel. Solar energy is currently adequately absorbed and stored in various ways and used in many applications such as solar water heating (Singh et al. 2021), building temperature maintenance (Faraj et al. 2020), solar distillation (Verma and Das 2021), solar pumping (Sivakumar et al. 2022), solar pond (Kumar and Das 2021; Verma and Das 2020), solar cooking (Sharma et al. 2005; Verma et al. 2022), solar power plant (Khatoon and Kim 2022), and solar drying applications (Srivastava et al. 2021; Verma and Das 2022). Initially, these methods of using solar energy had lesser efficiency. However, as research has progressed over the last decade, systems have improved and are now capable of replacing systems based on conventional fuel systems and being used effectively in various applications with better efficiency. Solar energy may currently be used efficiently in food drying applications also. Many industries involved in the food drying market are increasingly integrating their drying systems with solar drying systems in order to maximize profits.

Solar dryers have been created to alleviate the drawbacks of open sun drying. Food items are put in a thin layer for the drying process in the solar dryer's drying chamber. Solar radiation is employed to heat the drying chamber either by placing it in direct sunlight or by positioning it toward indirect sunlight. In a direct solar dryer, the drying chamber is covered with a glass sheet, trapping incident solar energy and evaporating moisture content from the food into the atmosphere. Solar collectors are used to elevating air temperature in an indirect solar dryer, and this heated air is cycled naturally or forcefully within the drying chamber, evaporating the moisture content of the food to the environment. Dried food quality in direct solar dryers is sometimes affected in terms of color and texture because of food that has been exposed directly to sunlight (Castillo Téllez et al. 2018). Solar dryers with the right design make solar energy use efficient and create high-quality dried foods. Solar drying is a non-carbon-emitting drying method. It's a superior solution for industrial food drying. It cuts down on the use of fossil fuels and lowers carbon emissions. Using a hybrid system for industrial food drying allows for continuous drying while decreasing the usage of traditional energy sources, lowering carbon emissions throughout the drying process, and maintaining the quality of dried products.

Infrared-aided drying, microwave-assisted drying, ultrasound-assisted drying, superheated steam drying, high electric field drying, and hybrid method are some of the food drying processes used in industrial drying systems (Khaing Hnin et al. 2019; Moses et al. 2014). To improve the drying performance of a certain crop, pretreatment is

also employed before drying in industrial drying systems (Llavata et al. 2020). Electrical heater-based drying systems and heat pump-based drying systems are also utilized in the industry for drying (Moses et al. 2014; Perera and Rahman 1997), but the cost of drying with these methods is greater and requires adequate system maintenance. These drying procedures also lose some of the energy used in the drying process, making them less effective (Menon et al. 2020). Because solar energy is free, combining a solar drying system with an industrial drying system is preferable to reduce drying costs (Rathore and Panwar 2011). When an industrial drying system is integrated with a solar drying system for drying food items, conventional fuel is saved, and carbon emissions are reduced (Menon et al. 2020). The use of solar drying in conjunction with an industrial drying system provides both direct and indirect benefits to the user.

Previously, several researchers assessed solar drying system performance in terms of capacity, drying time, energy consumption, and efficiency. Various authors used various approaches to perform techno-economic research on various solar drying systems. Fudholi et al. evaluated solar drying systems in Malaysia using air-based solar collectors. This evaluation included energy, exergy, environmental, and economic analyses. According to the study, solar drying saved 46% more drying time than open sun drying, with savings ranging from 13 to 86%. Thermal efficiency, drying system efficiency, and pick-up efficiency were rated at 39%, 27%, and 41%, respectively. The lowest and highest drying efficiencies were 6 and 52%, respectively.

Furthermore, the minimum and highest SPPs were 0.6 and 5 years, with an average of 2.6 years (Fudholi et al. 2015). Ekka and Palanisamy designed a mixed-mode solar dryer. Various agricultural goods such as jackfruit, red chili, turmeric, and ginger were dried using a designed mixed-mode solar dryer tested throughout the year. The mixed-mode solar dryer's drying temperature was kept 15–33 °C above the ambient air temperature. The net present value (NPV), payback duration, life cycle savings, and IRR were used to calculate the economic assessment. The designed mixed-mode dryer's SPP was calculated at 0.9 years. On the mixed-mode dryer initial capital investment, the IRR was 130%. On the basis of equivalent thermal energy, the environmental impact of the designed mixed-mode solar dryer was examined. The developed dryer has an energy payback time of 3.75 years and a net CO_2 reduction potential of 8.4 tons over its lifetime (Ekka and Palanisamy 2021). Hamid et al. did a technical and economic study of a hybrid type solar dryer using a collector of vacuum tube type. In order to determine the profitability and quantity of drying in this dryer, an economic study was performed. The results revealed that the SPP of this dryer was 3.7 years and that a total capacity of 1400 kg was preferable to open drying, with a yearly yield of 15,723 USD (Hamid et al. 2020). Nwakuba et al. developed a hybrid solar dryer integrated with an electric heater and evaluated its technical and economic analyses. The analysis indicated that with a low SPP of 0.72 years, this developed hybrid solar-electric dryer might save up to $1490.33 per year, making the dryer cost-effective and commercially feasible (Nwakuba et al. 2020). Many other technical and economic analysis was performed on different dryers that showed the overall benefits of the solar drying system (Fudholi et al. 2011; Selvanayaki and Sampathkumar 2017; Lamrani and Draoui 2021).

Various RETScreen models are developed to evaluate the performance, environmental and financial feasibility of various solar applications such as solar water heating (Singh et al. 2020), maintaining building temperature (Radosevic et al. 2022), solar PV systems (Asad et al. 2022; Ashofteh and Forough 2022), solar power plant (Windarta et al. 2022; Windarta et al. 2022) and other applications in order to provide a clear picture of the overall benefits of a solar system in a specific region. The industrial solar drying RETScreen model has yet to be investigated to determine the total benefits of this drying system to investors of India in the food drying business. The current study combines an industrial drying system with a solar drying system to keep the drying chamber at 40 °C and a 29.4 kg/s flow rate. This system's performance and net profitability were tested in six different locations across India. The results revealed the optimum location where this system provides high performance and net profitability, allowing industrialists to invest in the most profitable location with the best quality drying. The study discusses the many technological, environmental, and economic considerations of employing a solar drying system in the Indian climate.

Solar drying technology

In developing countries, the traditional drying method spreads the material in a fine layer at an open sun stage. However, apart from a few workers, natural drying does not require money; however, since there is no control over the drying rate, it is unhygienic, and the quality of products suffers. Solar energy can be used successfully for controlled drying, resulting in high-quality products (El Hage et al. 2018). One of the methods for preserving horticultural products is solar drying. According to researchers, drying is a dynamic heat and mass transfer process between the product and the drying environment. Solar drying research has shown that fast, modest, and effective solar dryers are needed to meet the increased food conservation demand, especially in developing countries like India. Many solar dryer systems have been studied, and excellent testing is still being done for multiple products (Lakhani and Patel 2013). Controlled drying, in general, involves controlling

drying parameters such as air temperature, humidity, drying rate, moisture content, and wind speed. Since there are so many alternatives in the solar dryer package, the design of solar dryers has much flexibility. This solar dryer is classified from a variety of perspectives. Dryers can be classified into two types based on basic design and abilities. These are direct or natural convection dryers and indirect type or forced circulation dryers (Pirasteh et al. 2014).

Direct solar dryer does not need a fan or a blower to deal with using electrical methods; such dryers are best suited for use in developing countries. These are both inexpensive and simple to build. Problem with such dryers is slow drying, less temperature and humidity control, a small volume that can be dried, and some damaged products due to direct sun exposure (Kumar et al. 2014), which may change the shade and flavors. They have a kind of enclosure and a translucent cover in their most basic shape. Because of the high temperature of the fenced-in region, the food products are heated by direct heat absorption, and the moisture in the substance dissipates and flows into the normal or natural air circulation (Tiwari 2016). Direct solar dryers come in various designs, which are manufactured as per the availability of local materials or the drying of a specific product to be dried (Lingayat et al. 2020). Some dryers are produced, checked, and evaluated in different countries. Schematic of direct solar dryers is shown in Fig. 1. Box style or cabinet dryer is the simplest example of a direct solar dryer. A direct solar dryer is widely used in places that receive direct sunshine for prolonged periods during the day. Cabinet is made of wood and is entirely insulated on the inside, behind, and bottom walls. A dense glass panel covers the slanted front wall, allowing sunshine to pass in (Hii et al. 2006). In this dryer, solar radiation is immediately absorbed by the food product put on the tray, reducing moisture content and allowing humid air to exit the dryer through a ventilation hole on the top of the dryer. Because the glass cover captures solar energy, the temperature inside the dryer is higher than outside. The main advantage of this dryer over open sun drying is that it protects food products from dust and foreign particle inclusion, and it also takes less time to dry since the temperature within the dryer is greater than the temperature outside. Its construction is also easier, and it may be made from locally available materials. However, the sole drawback of this dryer is that the dried product sometimes has poor quality in terms of color and aroma (Tiwari et al. 2013).

Direct solar dryers may affect the color, texture, and aroma of dried products due to direct exposure to solar radiation. However, it is better than open sun drying as the drying rate is faster and provides better quality than open sun drying.

The food product is not placed directly under solar radiation exposure in an indirect solar dryer. In these dryers, solar radiation is used to increase air temperature, and this warmed air is circulated within the drying chamber. Warmed air takes moisture content from foodstuff and escapes from the dryer by the natural circulation or forced circulation of hot air (Srivastava et al. 2021). Air movement in the solar dryer can be classified as natural convection indirect solar dryers and forced convection solar dryers. In the natural convection dryer, heated air circulated within the dryer naturally. In some cases, a chimney is installed in the solar dryer to improve air velocity within the dryer without using a fan/blower (Vijayan et al. 2020b).

In forced convection solar dryers, fans or blowers circulate heated air within the dryer. Air velocity can be controlled in these solar dryers (Hajar et al. 2017). Figure 2 shows a natural convection solar dryer with a chimney, also known as a passive solar dryer. Solar flat plate collectors convert solar radiation into heat energy in the air (Vijayan et al. 2020a). The drying chamber is directly connected to a solar collector, and the dryer has a chimney on top. A black-painted copper sheet is placed bottom of the solar collector. A clear glass sheet with some air holes is mounted on the top of this copper sheet to make the airflow within

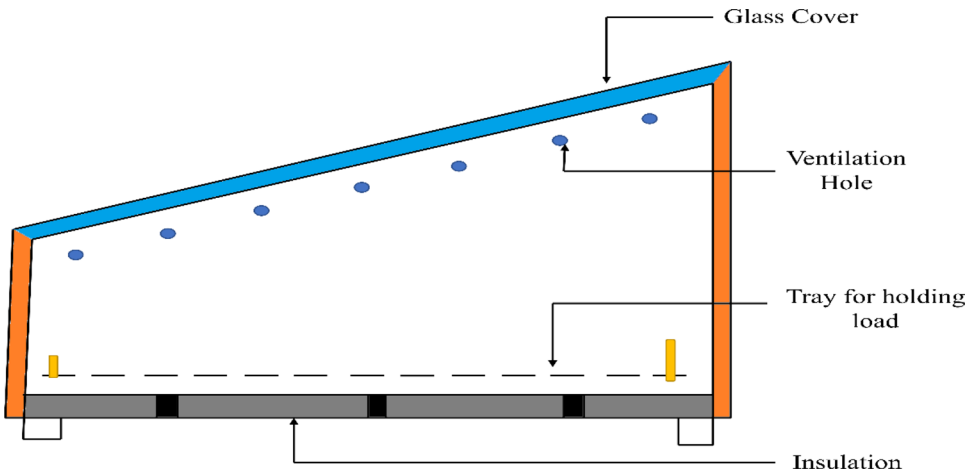


Fig. 1 Cabinet-type solar dryer

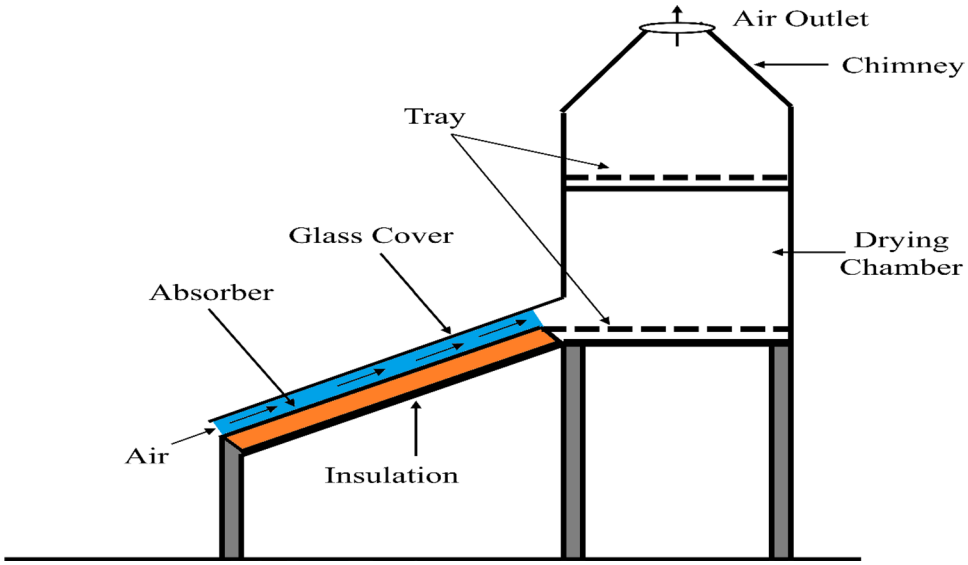


Fig. 2 Indirect solar dryer with chimney

the dryer. To minimize heat losses, the collector's base is insulated from the collector plate's base surface. One end of the collector is exposed to environment, while other is attached to the foundation of the wooden drying chamber. The room's floor is painted with dark black shading to maximize the overall obtained heat. This room's chimney is built on the roof, providing a natural forced effect on the airflow. There is a pressure differential inside the chamber due to the system's design, which allows for wind movement (Vijayan et al. 2016). As a result, the pressure differential, which usually induces natural circulation, is present in and out of the dryer. Food products to be dried may be put on a tray made of a non-corrosive metal plate in a chamber (Lingayat et al. 2017).

This research uses an indirect forced circulation solar dryer with an electrical heater for an industrial food drying facility. Previously hybrid drying systems were used for drying application in which solar drying system is integrated with some other source of energy to facilitate solar drying system and improves the overall performance of the hybrid drying system (Kamarulzaman et al. 2021). Because energy is required in many other applications in industry, it is best to integrate solar dryers with electric heaters in industrial solar drying applications. Using an electric heater as an auxiliary source in solar drying would seldom disrupt the drying process (Koçak et al. 2020). Previously, solar dryers integrated with electric heaters were researched, and their performance was evaluated, but not in accordance with the needs of commercial food drying applications (Wang et al. 2018; Amer et al. 2018; Lamrani et al. 2019; Lamrani and Draoui 2020). In the current study, indirect forced circulated solar dryers with an electric heater for industrial application are analyzed, enlisting actual and overall benefits of this system. Air forcedly circulated at 29.4 kg/s within the drying chamber in the current system. Dryer has a 100 m^2 solar collector unit that absorbs solar energy and heats the air. For forced air circulation, a fan is utilized. Auxiliary electrical heaters are connected between the fan and the drying chamber, ensuring that the air entering the drying chamber is at least 40 °C. Figure 3 shows a schematic representation of the industrial solar dryer used in the current study. Because this system is deployed in diverse locations across India, the findings of this study demonstrate the locations where this industrial drying system provides the best outcomes. Research' findings encourage industrialists to invest in and establish a food industry in that region, ensuring that investors get the most out of this drying system.

Implementation of RETScreen expert software

RETScreen expert is an energy modeling tool that offers a comprehensive and effective platform for assessing the viability of different types of renewable energy models. It is a useful method for those who decide whether to support or oppose a proposal during the early stages of growth. These packages provide tools for assessing any kind of project, such as a tracker of performance, interactive energy analyzer, financial risk assessor, and smart project identifier. Model 1 of RETScreen is used in the current study of the solar dryer project to do a feasibility study on a solar dryer system in several cities throughout India. Energy model includes six worksheets [Energy Modelling, Solar Resource and Calculating the Heating Load, Sensitivity and Risk, Cost Analysis, Greenhouse Gases (GHG) Emissions Reduction calculations, Financial Summary, Sensitivity, and Analysis of Risk (Natural Resources Canada 2005).

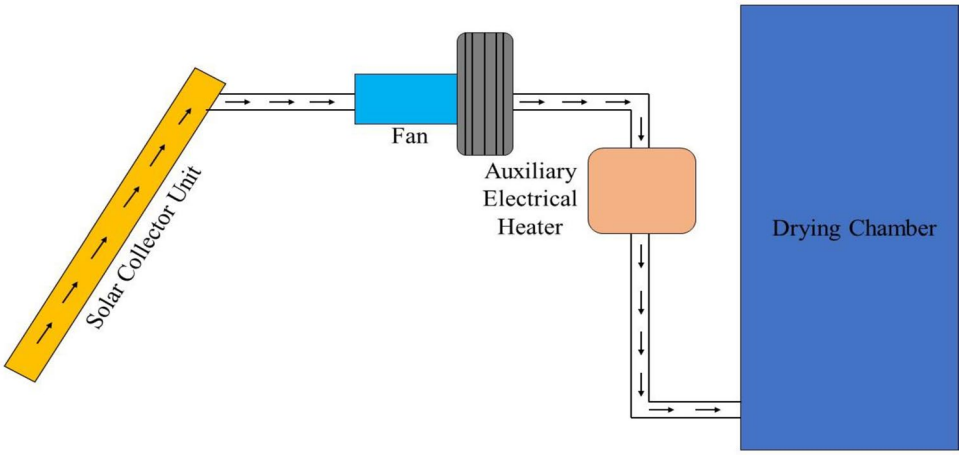


Fig. 3 Industrial solar dryer used in the current study

Solar Resource (SR) and Heating Load Calculation (HLC) worksheet calculate the monthly energy requirement needed to heat the drying chamber to a required temperature. This worksheet uses monthly solar radiation estimates on a horizontal surface to measure the annual solar radiation on the tilted collector array on every array orientation. Annual efficiency of a solar drying system is determined by the drying system's characteristics, the availability of solar energy, the temperature of the surrounding air, and the characteristics of the heating load (Kushwah et al. 2021). Energy model calculates the project's cost determined by input parameters such as drying chamber capacity/solar collector area, number of collectors, and heat exchanger. GHG analysis worksheet is presented to quantify the planned project's GHG reduction capacity. The sensitivity worksheet determines the sensitivity of critical financial metrics to key technological and financial constraints. The energy model worksheet and the SR & HLC worksheet were completed first. Following that, the cost estimation worksheet should be done, followed by the financial summary worksheet. Software RETScreen is a one-of-a-kind decision-making platform for tracking, reviewing, and assessing the operating efficiency of renewable energy systems. Current study base case is considered to run on diesel fuel to maintain 40 °C and an airflow rate of 29.4 kg/s and the proposed case is running on electricity and solar energy with maintaining the same output conditions for drying operations.

City selection criteria

Six cities from different geographical zones of India have been chosen to illustrate the advantages of the solar dryer system at different locations. These six cities are selected to demonstrate different environmental conditions on the solar dryer. Table 1 shows the climate data locations for all of the cities that were chosen.

Hanumangarh, Rajasthan

Hanumangarh, Rajasthan is located at Latitude 29.6° N and Longitude 74.3° E and comes under extremely hot and dry conditions. It receives about 4.89 kWh/m^2/day of excellent solar radiation. In Hanumangarh the temperature rises to 24–25 °C and drops at night to 14–15 °C, without precipitations. Variation of solar radiation value throughout the year in Hanumangarh is shown in Fig. 4. It is a mainly a desert region with little rainfall. Major farm activities include cultivation of rice, millet, cotton, sonamukhi (senna), wheat, and vegetables. For the drying of crops, the solar dryer system may be utilized. Excellent available solar radiation

Table 1 Selected cities of India for solar drying system modelling. *Source*: NASA database

Cities	Latitude (° N)	Longitude (° E)	Elevation (m)
Hanumangarh, Rajasthan	29.6	74.3	185
Bangalore, Karnataka	13.0	77.6	921
New Delhi, Delhi	28.6	77.1	236
Lucknow, Uttar Pradesh	26.8	80.9	128
Shimla, Himachal Pradesh	31.1	77.2	1770
Tiruchchirappalli, Tamil Nadu	10.8	78.7	88

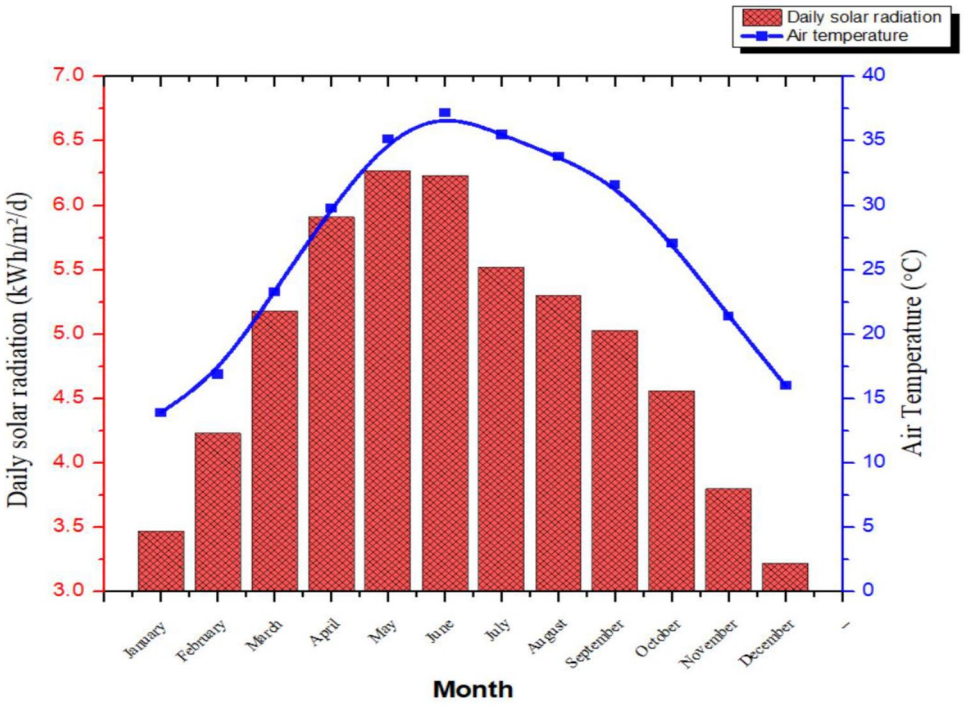


Fig. 4 Climate data of Hanumangarh. *Source*: National Aeronautics and Space Administration (NASA) database

and dry climate condition promote solar drying systems in Hanumangarh.

Bangalore, Karnataka

Bangalore is located at 13.0° N & 77.6° E and comes under hot and humid conditions. The city has a moderate climate condition and receives nearly 5.32 kWh/m^2/day solar

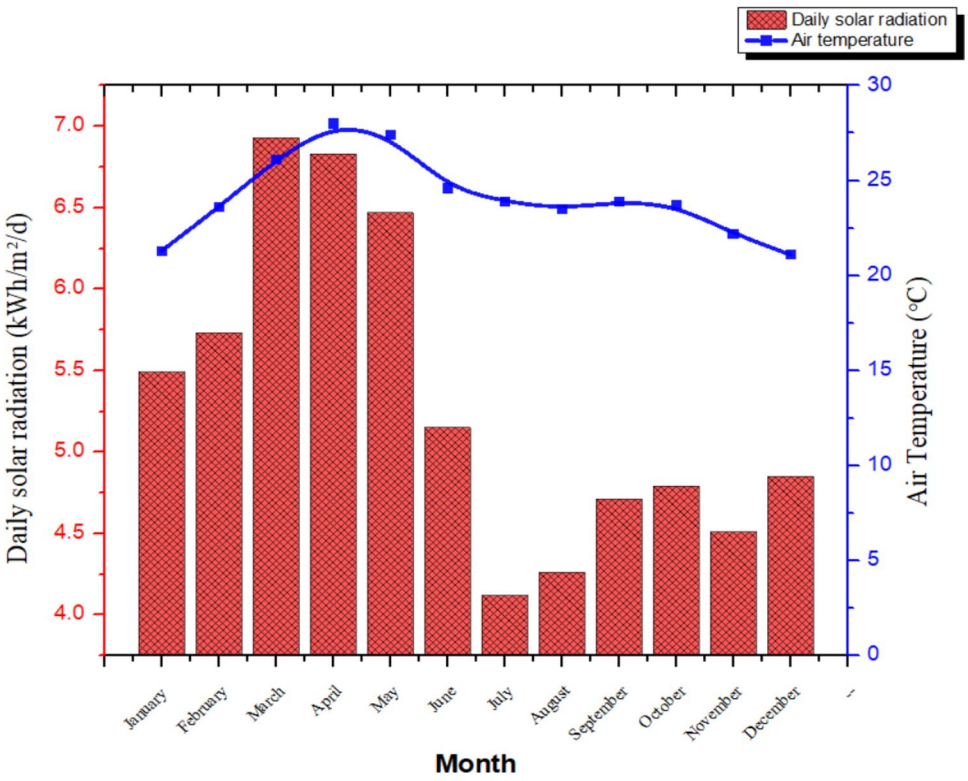


Fig. 5 Climate data of Bangalore. *Source*: NASA database

radiation. Variation of solar radiation value throughout the year in Bangalore is shown in Fig. 5. This city is endowed with horticultural and agricultural crops such as ragi, sugarcane, groundnut, rice, castor, mulberry, grapes, etc. A solar dryer system may be used for crop drying with maintaining an excellent quality of the dried product. Utilization of the solar drying system in Bangalore promotes excellent, accessible solar radiation and dry conditions.

New Delhi, Delhi

New Delhi is located at 26.8° N & 77.1° E and has a mixed climate. It receives a lot of solar radiation, around 5.25 kWh/m^2/day. Variation of solar radiation value and temperature throughout the year in New Delhi is shown in Fig. 6. Cauliflower, spinach, carrot, cabbage, mustard, tomato, and okra are cultivated in and around Delhi. Wheat, paddy, bajra, and jawar are the most important crops cultivated in Delhi. Solar drying systems can be used to dry vegetables and crops. Excellent solar radiation availability and the composite climate of Delhi promote solar drying systems to produce high-quality dried products.

Lucknow, Uttar Pradesh

Lucknow is located at 28.6° N & 80.9° E and falls within mixed climate conditions. It receives a significant amount of solar radiation, around 4.96 kWh/m^2/day. Variation of solar radiation value and temperature throughout the year in Lucknow is shown in Fig. 7. The main crops grown in Lucknow are wheat, rice, Mentha, field pea, mustard, sugarcane, etc. vegetables and fruits grown in and around Lucknow include potato, brinjal, okra, vegetable pea, cabbage, mango, banana, guava, papaya, etc. During the harvesting season of the product, a solar drying system can be used to dry the food product. High solar radiation capacities in Lucknow encourage the usage of solar drying systems to produce high-quality dried products to reduce the chance of food loss during the harvesting season.

Shimla, Himachal Pradesh

Shimla is located at 31.1° N & 77.2° E. It has a cold climate condition and receives a lot of solar radiation, nearly 5.28 kWh/m^2/day. Variation of solar radiation value and

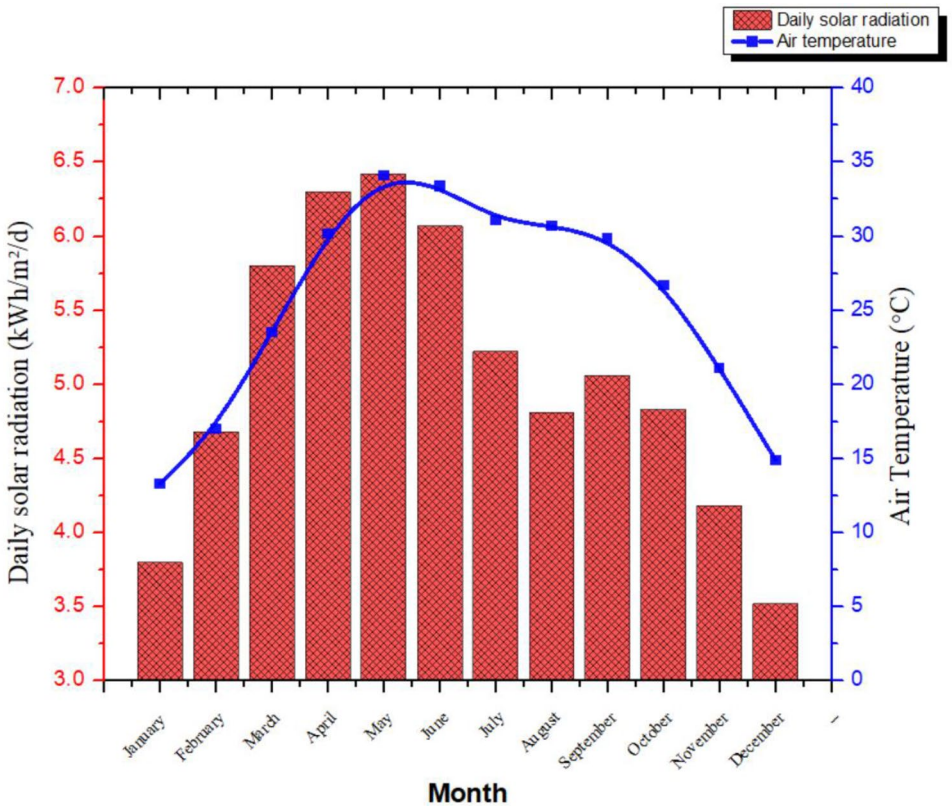


Fig. 6 Climate data of New Delhi. *Source*: NASA database

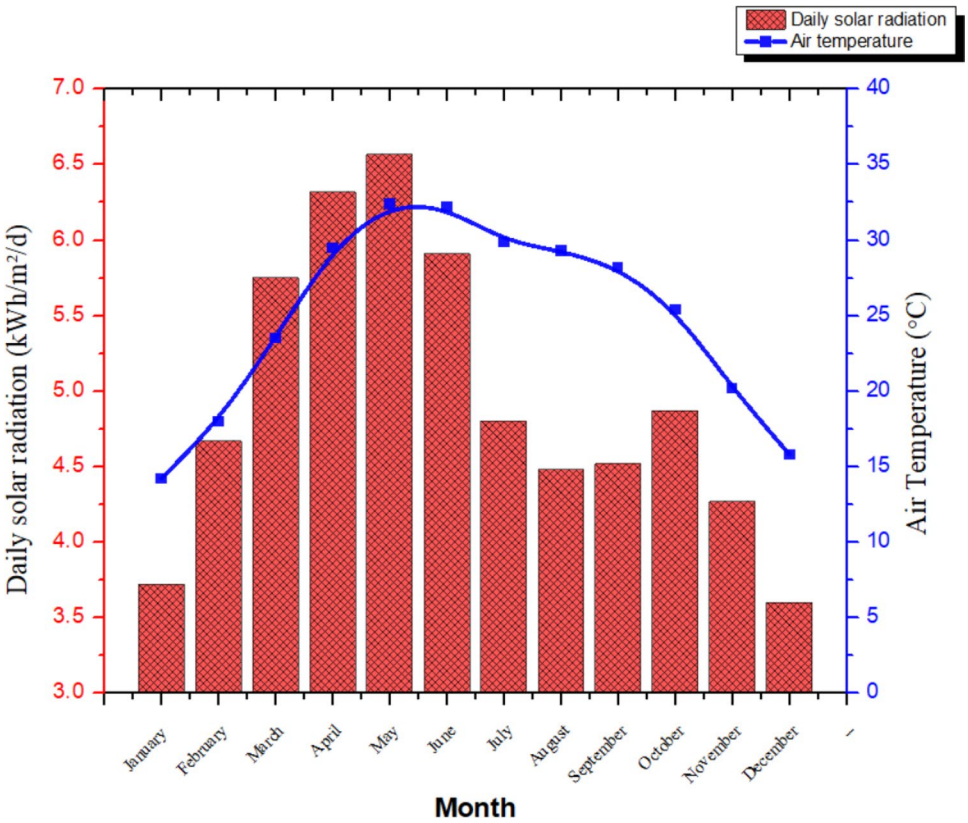


Fig. 7 Climate data of Lucknow. *Source*: NASA database

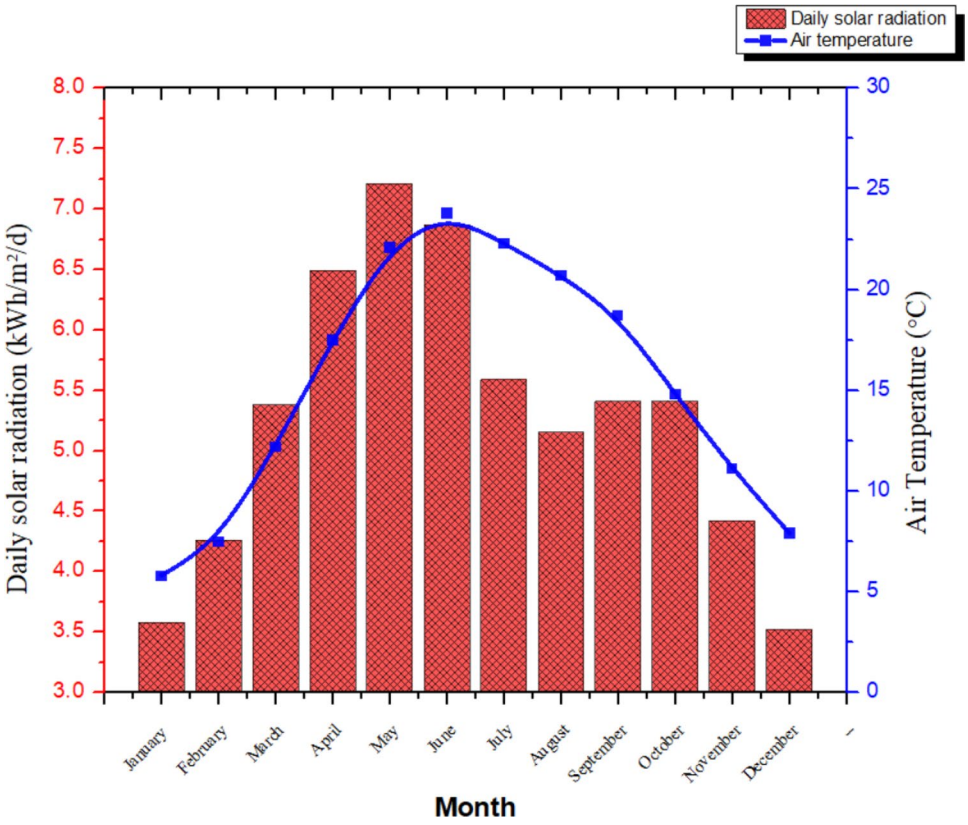


Fig. 8 Climate data of Shimla. *Source*: NASA database

temperature throughout the year in Shimla is shown in Fig. 8. Fruits grown in and around Shimla include apple, peach, almond, apricot, plum, pear, pecan nut, pomegranate, walnut, strawberry, kiwi, etc. major field crop cultivated are maize, paddy, wheat, barley, pulses, and oilseeds. In Shimla, a solar drying system can be utilized to dry food. The high-quality dry fruits of Shimla have huge demand and the application of solar drying systems give a high-quality dried product. During foggy season auxiliary sources can be used in conjunction with a solar dryer to dry food products.

Tiruchchirappalli, Tamil Nadu

Tiruchchirappalli is located at 10.8° N & 78.7° E latitude and longitude. It has a hot and humid climate and gets much solar radiation, nearly 5.28 kWh/m^2/day. Variation of solar radiation value and temperature throughout the year in Tiruchirappalli is shown in Fig. 9. Tiruchirappalli crops includes fruits (37%), vegetables (34%), plantation (20%), spices (4%), flowers (3%), and medicinal crops (3%). banana, coconut, tapioca, onion, mango, acid lime, and chilies are the major horticulture crops that are cultivated in the district. Solar drying systems can be employed for drying these crops. Excellent solar radiation availability and hot climate conditions can promote the use of solar drying systems for the drying of excellent quality dried products.

Metrological data of selected six cities

In this work, metrological data such as wind speed, air temperature, solar radiation, and so on are gathered from NASA's database for simulating the solar drying system. Table 2 contains metrological statistics for all six cities of India. Daily solar radiation on a horizontal surface is the most important climatic element for the drying process in the solar drying system. Figure 10 depicts the monthly fluctuation of daily solar radiation over the horizontal surface of all six cities. Relative humidity and wind speed play an important role in solar drying. The monthly data for the relative humidity of all selected six cities are shown in Fig. 11.

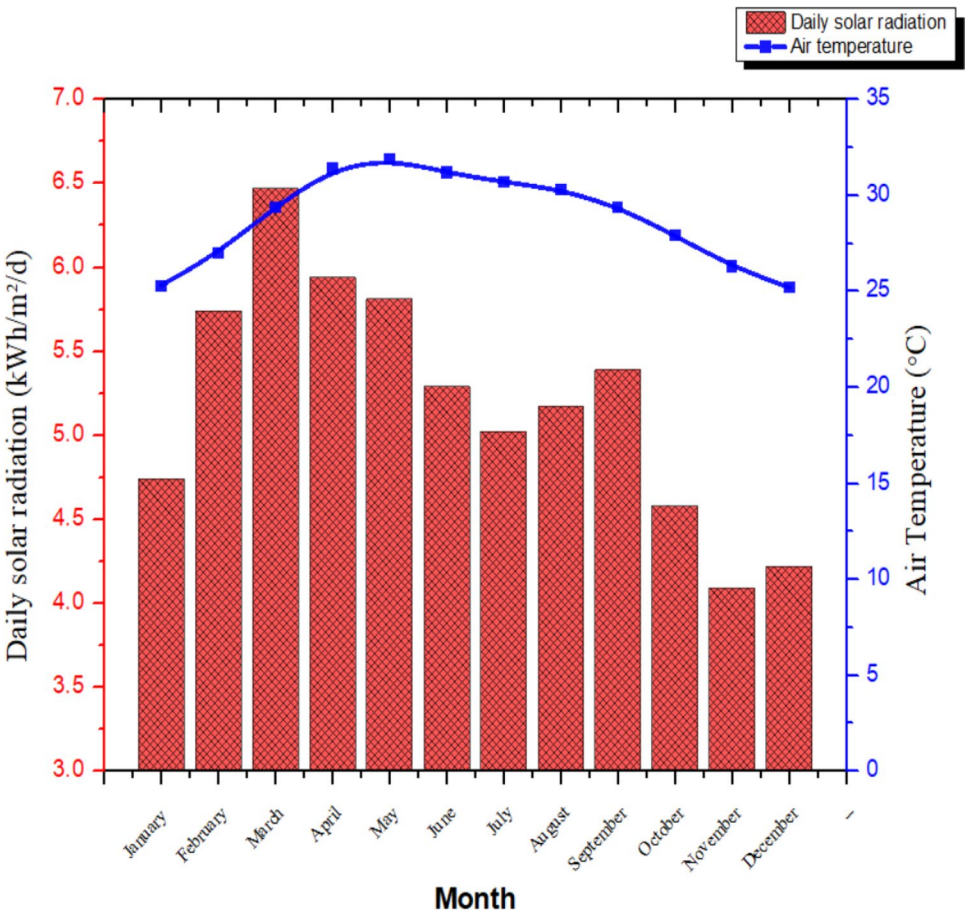


Fig. 9 Climate data of Tiruchchirappalli. *Source*: NASA database

Table 2 All selected cities' annual metrological data (NASA database)

Metrological parameter	Annual values					
	Hanumangarh	Bangalore	New Delhi	Lucknow	Shimla	Tiruchchirappalli
Air temperature (°C)	26.9	24.1	25.5	24.9	15.4	28.8
Relative humidity (%)	31.5%	68.6%	58.1%	67.1%	49.9%	66.8%
Precipitation (mm)	325.99	852.14	710.82	827.08	949.90	800.59
Solar radiation horizontal on daily basis (kWh/m^2/d)	4.89	5.32	5.06	4.96	5.28	5.20
Atmospheric pressure (kPa)	98.6	92.6	98.2	99.3	82.4	99.4
Wind speed at 10 m (m/s)	2.8	1.9	2.5	2.0	2.4	4.5
Earth temperature at 0 m (°C)	27.3	25.1	25.9	26.5	14.6	28.8
Heating degree days (at 18 °C)	220	0	270	186	1486	0
Cooling degree days (at 10 °C)	6152	5149	5667	5443	2238	6877

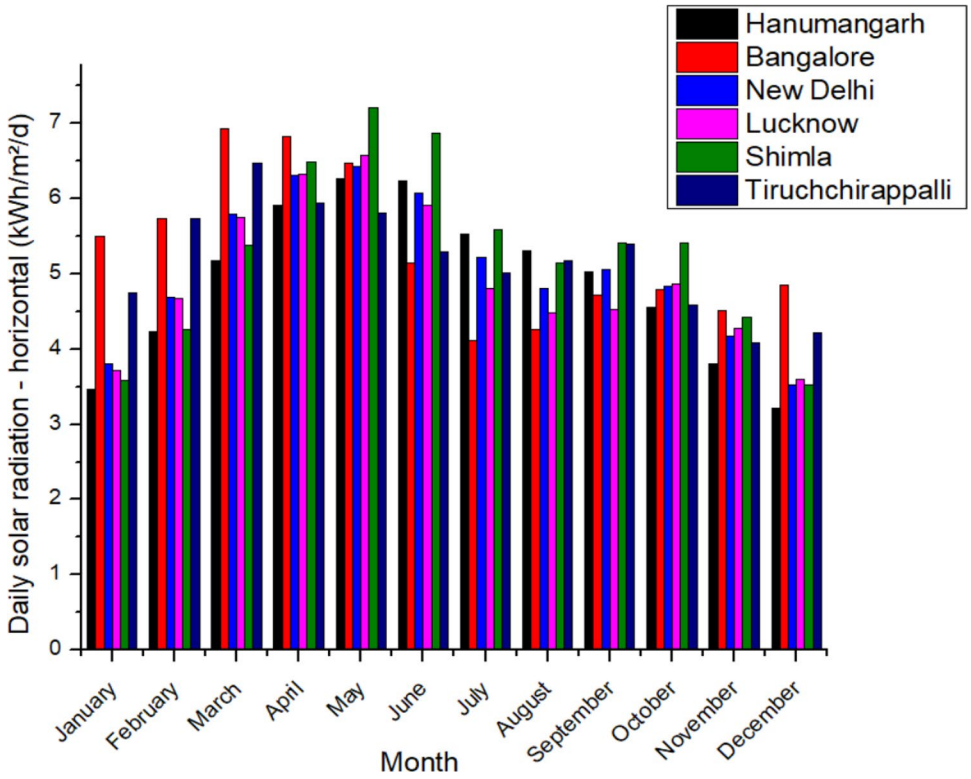


Fig. 10 All chosen cities' daily solar radiation

Techno-economic input parameters for simulation

The term "technical–economic" relates to determining the best place for installing a solar drying system while spending the least amount of money on the project. Simulation model allows for identifying project locations with related costs. This sort of research is critical for determining the feasibility of a project in a geographically diverse region anywhere in the world. Terminology techno refers to the technical details of the components, such as the kind of solar air heater and the manufacturer. Economic refers to financial indicators such as EPP, SSP, IRR, and modified internal rate of return (MIRR) (Singh et al. 2020). A year to positive cash flow is another word for the equity payback period (EPP). This is the first year that the project's cumulative cash flows are positive and T_{PCF} denotes it. The following equation for T_{PCF} is used to compute it (Singh et al. 2020).

$$0 = \sum_{t=0}^{T_{PCF}} C_t \tag{1}$$

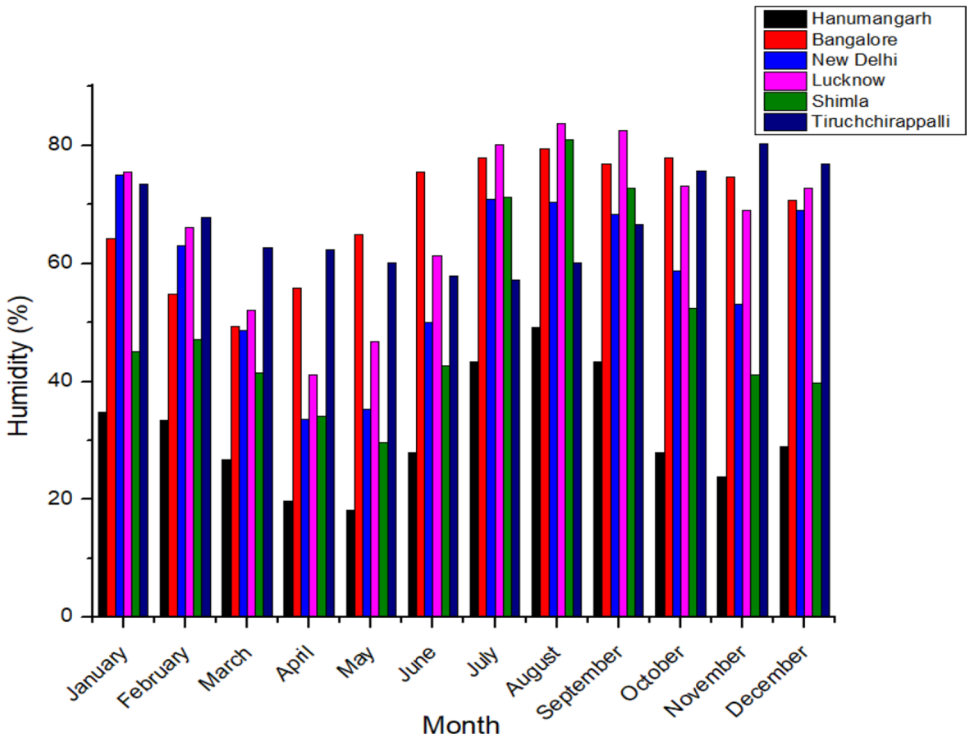


Fig. 11 Monthly relative humidity value of selected cities

Table 3 RETScreen input parameters for the Energy Model for the Basic Scenario Model 1

Input parameters	Value
Indoor temperature (°C)	21
Air temperature maximum (°C)	40
Design air flow rate (L/S)	24,000
Solar tracking mode	Fixed
Azimuth	180
Slope	22
Operating days per week (d/w)	7
Operating hour per day (h/d)	12
Type of Solar Air heater	Transpired-plate
Miscellaneous losses (%)	3
Type of fuel	Electricity
Seasonal efficiency (%)	54.2
Annual rate of electricity (USD/kWh)	0.11
Cost escalation rate of fuel (%)	1.5
Rate of discount (%)	9
Rate of reinvestment (%)	9
Rate of inflation (%)	4
Life of project (years)	30
Debt ratio (% of the initial cost)	75
Debt term	15
Cost of annual O&M (USD)	0
Rate of GHG reduction credit (USD/tCO_2)	20

Table 4 Technical specification of the selected solar air heater

Type	Transpired- plate
Objective of design	High air volume
Manufacturer name	Conserval engineering
Model name	Solar wall
Absorptivity of solar collector	0.95
Performance factor	1
Area of solar collector	100 m^2
Capacity	70 kW

Simple payback period (SPP) can be calculated by the following relation, which shows the number of years that the project takes to return the initial investment (Singh et al. 2020). SPP is simply the ratio of actual initial investment to net cash flow per year

$$\mathrm{SPP} = \frac{C - \mathrm{IG}}{\left(C_{\mathrm{E}} + C_{\mathrm{C}} + C_{\mathrm{RC}} + C_{\mathrm{GHG}}\right) - \left(C_{\mathrm{O\&M}} + C_{\mathrm{fuel}}\right)} \tag{2}$$

Internal rate of return (IRR) is a financial statistic used to calculate the profitability of possible investments. IRR can be calculated from the following relations

$$\mathrm{NPV} = \sum_{t=0}^{T} \frac{C_t}{(1 + \mathrm{IRR})^t} = 0 \tag{3}$$

Table 3 shows the basic scenario model 1 input parameters. The current study considered a solar air dryer of 100 m^2 in each of the selected six cities to model the solar drying system within India. At first, no grant for purchasing a solar drying system is taken, and the grant's effect is then analyzed in terms of SPP. Table 4 shows the specifications of the selected solar air heater for simulation. The simulation method is compared with a base case considering the same inputs running of diesel fuel and the proposed case is running with a combination of electrical and solar energy. In the proposed case 100 m^2 solar air heater is integrated with an electrical system to fulfill the energy requirement of the dryer.

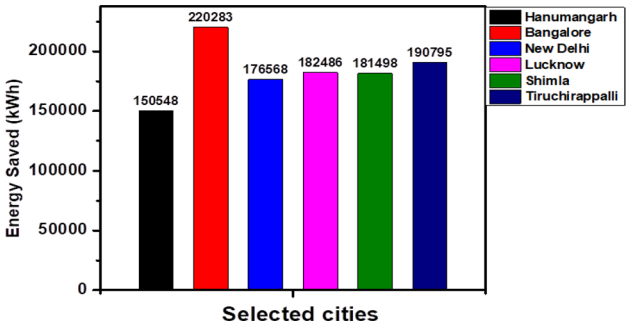


Fig. 12 Solar drying system's yearly potential for energy savings

Results and discussion

The output parameters of all cities from the basic scenario model 1 are summarized in Table 5. In all cities, the specifications of solar air heaters are the same. The proposed case is compared to the base case running on diesel for heating purposes. The base case is integrated with a solar collector area of 100 m^2 and the results of the simulation are tabulated. In Bangalore, maximum energy is saved and the gross annual GHG reduction is higher.

Energy savings

The performance of the drying system of the proposed case is evaluated by energy savings using a 100 m^2 area solar air heater compared with the base case running on diesel oil. Energy requirement for maintaining temperature and airflow rate is fulfilled by electricity and solar energy in the proposed case. Energy saving of all selected six cities is shown in Fig. 12. The simulation result of the proposed case shows that in Bangalore maximum energy is saved amounts 220,283 kWh by using a 100 m^2 solar air heater followed by Tiruchchirappalli (190,795 kWh) and Lucknow (182,486 kWh). Hanumangarh shows minimum energy savings of amount 150,548 kWh. Energy savings data also indicates that the solar fraction is highest in Bangalore followed by Tiruchchirappalli, Lucknow, Shimla, New Delhi, and Hanumangarh. Because of the varying amounts of solar fraction in different cities, there is a difference in yearly energy savings. Bangalore saves the most energy for the present model since solar radiation is practically constant throughout the year and the humidity in the environment is higher than in other

Table 5 Output parameters of selected cities from the basic scenario model 1

Output parameters	Values					
	Hanumangarh	Bangalore	New Delhi	Lucknow	Shimla	Tiruchchirappalli
Maximum air temperature (°C)	40	40	40	40	40	40
Design air flow rate (kg/s)	29.4	29.4	29.4	29.4	29.4	29.4
Operating hours per day (H/d)	12	12	12	12	12	12
Solar collector area (m^2)	100	100	100	100	100	100
Capacity (kW)	70	70	70	70	70	70
Initial cost (USD/m^2)	270	270	270	270	270	270
Energy saved (kWh)	150,548	220,283	176,568	182,486	181,498	190,795
Gross annual GHG reduction (t CO_2)	38.1	55.7	44.6	46.1	45.9	48.2
Gross GHG reduction (t CO_2)-30 years	1143	1671	1338	1383	1377	1446
Equivalent car and light trucks not used	7	10.2	8.2	8.4	8.4	9
Simple payback period (years)	1.9	1.3	1.6	1.6	1.6	1.5
Pre-tax IRR equity (%)	183	285	221	230	228	242
Pre-tax IRR assets (%)	49.2	74.7	58.7	60.9	60.5	63.9
Equity payback period (years)	0.56	0.36	0.46	0.44	0.45	0.42

cities. Solar fraction value can be enhanced by integrating a large area of a solar air heater in the proposed case.

Reduction of greenhouse gas emissions

Solar drying does not produce any greenhouse gases. By adopting renewable energy technology will be used to fulfill the hot air requirement for drying operations, the simulation model determines a certain amount of GHG emission reduction. The potential for GHG emission reduction is a significant environmental occurrence that demonstrates the value of solar drying system projects in mitigating the effects of global warming. Simulation model allows users to compute GHG emissions every year as well as for the life of the project. It is determined in two ways: the gross yearly greenhouse gas (GHG) emission reduction in tCO_2/year and the gross greenhouse gas (GHG) emission reduction over the project's whole life, that is, tCO_2-30 year if the project's lifespan is set to 30 years. Gross annual GHG emission reduction indicates the number of possible emission reductions for a specific year, whereas the gross GHG emission reduction shows the total value over the project's lifetime. This study computed the solar drying system's gross yearly GHG emission reduction potential for all selected cities and displayed in Fig. 13.

Gross annual GHG emissions reduction potential of solar drying systems ranges from 38.1 tCO_2 to 55.7 tCO_2 (Fig. 13). The least value is obtained for Hanumangarh (38.1 tCO_2) and the highest value for Bangalore (55.7 tCO_2). The highest value is followed by two cities, namely Tiruchchirappalli (48.2 tCO_2) and Lucknow (46.1 tCO_2). The following greatest value is attained for Shimla (45.9 tCO_2). The other value is for New Delhi (44.6 tCO_2). The results reveal that the modeled setup in Bangalore has the greatest capacity to reduce GHG emissions since energy savings in Bangalore

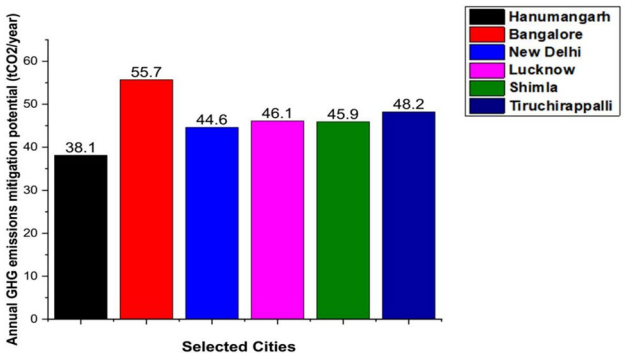


Fig. 13 Annual GHG emissions mitigation potential (tCO_2/ year)

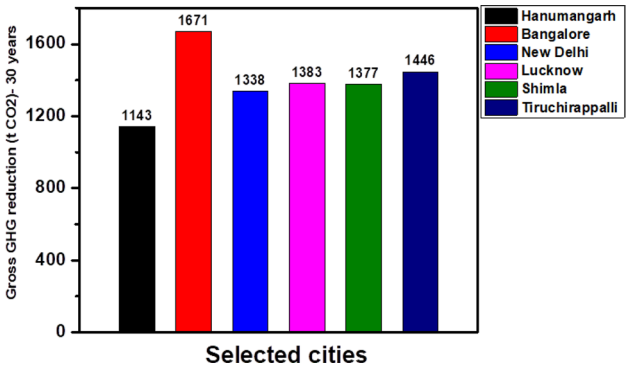


Fig. 14 Reduction in gross greenhouse gas (GHG) emissions (30 years) in the selected cities

is the highest, hence GHG emission reduction is likewise the highest among the selected cities. Figure 14 depicts the variance in gross GHG emission reduction in tCO_2 during the project period (30 years) for all chosen cities.

It has been found that the total decrease in gross GHG emissions for all cities is about 8358 tCO_2 in 30 years. Emission simulation model analyses the total yearly GHG emission reduction in terms of the number of vehicles and light trucks that aren't operating, which is a key environmental factor for evaluating solar drying system projects. The number of cars and light trucks not running in all six cities is given in Fig. 15. The graph clearly shows that Bangalore and Hanumangarh have the highest and lowest values. The second-largest value is obtained for Tiruchchirappalli and next for Lucknow. The difference in the number of automobiles and light trucks that are not operating is due to variations in energy savings utilizing the solar drying system, which is employed throughout the year based on the needs of a specific geographical region.

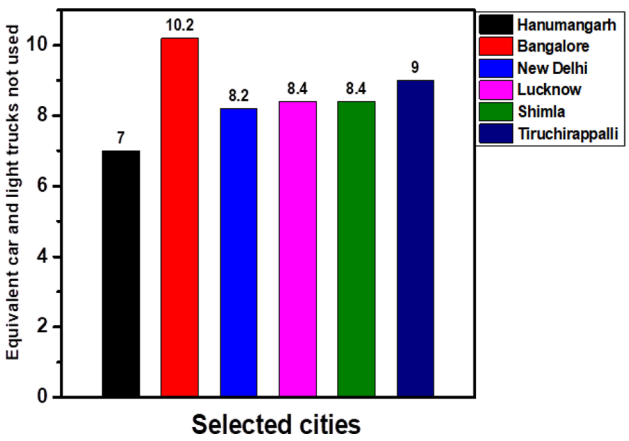


Fig. 15 An equivalent number of cars and light trucks are not used in selected cities

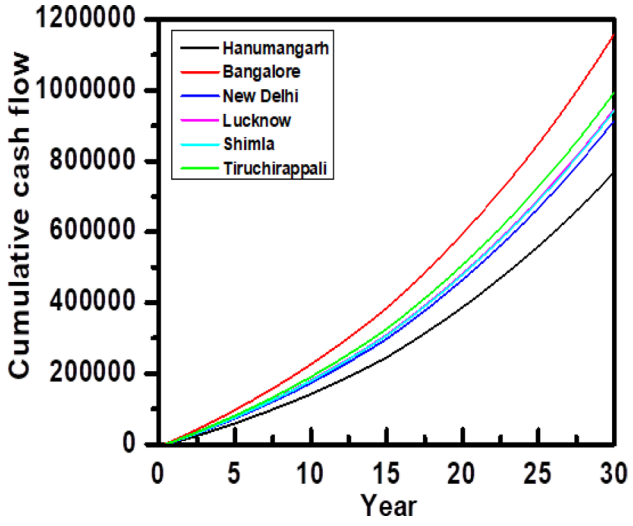


Fig. 16 Selected six cities' cumulative cash flow for project life

Financial analysis

The cost-effectiveness of installing solar drying system projects is assessed in this study using terminologies such as Equity Payback Period (EPP), Simple Payback Period (SPP), and Internal Rate of Return (IRR) equity. SPP is derived using graphs of progressive cash flow from the basic scenario model 1 with a debt ratio of 75%, a debt period of 15 years, and a debt interest rate of 9% (Singh et al. 2021). Figure 16 depicts the cumulative cash flow graph for selected cities. Initially, cumulative cash flow for all selected cities is negative for the first 1–2 years, but beyond that time, cumulative cash flow for all selected cities is positive. Bangalore has the highest cumulative cash flow slope, indicating that this system generates far more cash in the city using modeled solar drying system. Tiruchirappalli has the second-highest cumulative cash flow slope during the system's lifetime. This suggests that Tiruchirappalli generates more revenue than other cities after Bangalore.

SPP reflects the time it takes to repay its original cost through income or savings. The simulation model also computes EPP, which evaluates the amount of time required for a facility's owner to repay its preliminary investment (equity) via the project's cash flow. It indicates how fast the expense of an investment may be recouped. In the instance of executing a solar dryer project, for example, a negative unfavorable payback period would indicate that the year expenditures paid are more than the year savings gained. EPP incorporates the cash flow of the project from conception in addition to the project's leverage (debt amount), making it a stronger temporal indication of the project's benefits than the basic payback. Figure 17 depicts the SPP and EPP of all six cities for implementing solar drying systems. The figure shows that the largest value of EPP and SPP is found in Hanumangarh i.e., 0.56 years and 1.9 years respectively, while the minimum value of EPP and SPP is found in Bangalore i.e., 0.36 years and 1.3 years, respectively. This is because the solar fraction in Bangalore is the highest and the solar fraction in Hanumangarh is the lowest. Table 5 shows further values of output parameters.

IRR is the actual interest return produced by an investment in a project before taxation during the project's life. It is computed using pre-tax annual cash flows and the life duration of the project. IRR of all cities chosen for solar drying system installation is depicted in Fig. 18. The simulation result shows that the highest IRR value is found in Bangalore (74.7%), while the lowest is found in Hanumangarh (49.2%). This is because Hanumangarh has the lowest energy savings using the solar drying System, whereas Bangalore has the highest energy savings. Other cities like New

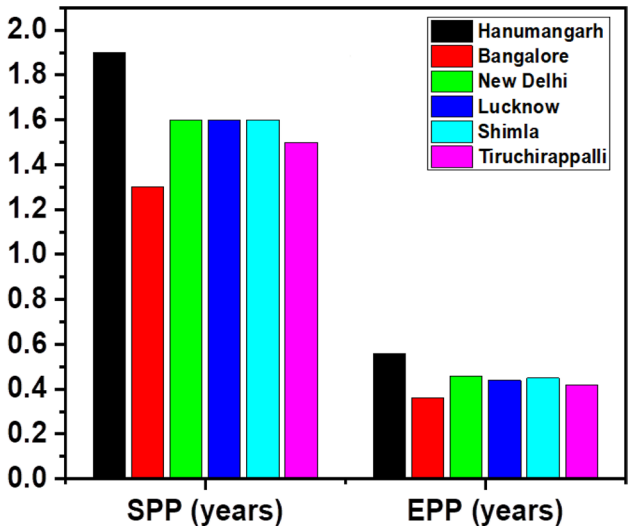


Fig. 17 Payback period of selected cities

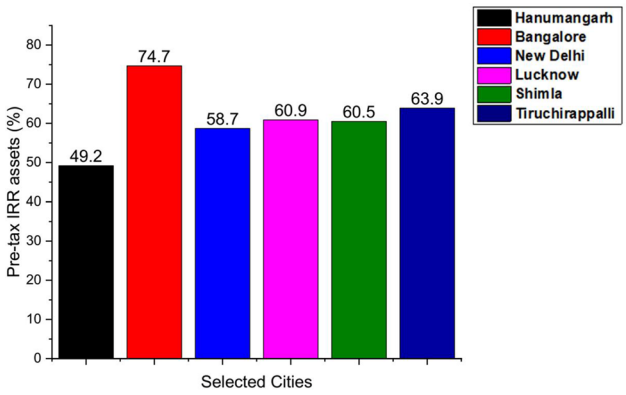


Fig. 18 All of the six cities' internal rates of return

Delhi, Lucknow, Shimla, and Tiruchirappalli have pre-tax IRR assets of 58.7%, 60.9%, 60.5%, and 63.9%, respectively.

Clean development mechanism potential (CDM)

CDM was established because of the Kyoto Protocol (1997) to make it simpler for wealthy countries to purchase carbon credits from renewable energy projects in developing nations. The necessity for this agreement is seen as arising from the fact that it costs more for affluent countries to preserve carbon credits in their respective nations than it does for developing countries. CDM executive board established some guidelines for project qualification, such as (a) energy optimization project activities that consume less energy by 60 GWh per year, (b) Projects involving renewable energy with a maximum production capacity of up to 15 MW are eligible, and (c) remaining project operations with an annual decrement in emissions a value lower than 60 $ktCO_2$. According to the current study, solar drying systems in India are excellent CDM approval applicants.

The average percentage uncertainty between the modeling findings and the actual energy yield RETscreen application was less than 5%, which is acceptable for feasibility and simulation studies (Lee et al. 2012a). Various uncertainties exist that lead to the findings deviating from the real energy production (Lee et al. 2012a, b). The basic mathematical procedures utilized in the modeling tools contain the initial uncertainty. These include, among other things, estimates of radiation on the array's plane extrapolated from global horizontal radiation, calculations of self-shading and horizon shading, and computations of low light levels. Low light levels and short-term high radiation peaks have been demonstrated in past research to be difficult for model algorithms to manage (EL-Shimy 2009; Owolabi et al. 2019). Other contributing factors to consider are variations in initial cost at different locations, debt ratio, debt interest rate, debt duration, and variations in GHG emission reduction credit rate (RETScreen® Software Online User Manual).

Conclusion and policy implications

Integrating a solar drying system without interfering with the drying process is beneficial for energy savings and lowering carbon emissions to the environment in the food drying industry, before installing the commercial solar drying system in a certain place. It is preferable to use the RETScreen simulation model to test the system's technological, environmental, and financial feasibility in various locations. According to the simulation results, the actual model should be positioned at the site where the greatest advantages are created. The results of this study give a clear picture to investors to optimize their profits by integrating solar drying systems into industrial food drying systems. The current work models a hybrid solar drying system to meet the needs of an industrial drying system. The RETScreen simulation model is intended to assess the system's technological, environmental, and financial feasibility. This study conducted a technical and financial feasibility analysis of the indirect solar drying system with electric heaters as per the requirement of the industrial drying system. Six cities from various geographical regions were chosen to determine the solar drying system's financial feasibility and performance. Bangalore has the highest daily solar radiation (5.32 kWh/m^2/day), while Hanumangarh has the lowest (4.89 kWh/m^2/day). As per the analysis results, Bangalore is the most suitable place for installing a solar drying system. It has also been found that all cities receive larger than 4.5 kWh/m^2/day of daily solar radiation, which is an excellent precursor for installing solar drying systems. Energy savings are highest in Bangalore (220,283 kWh) followed by Tiruchchirappalli (190,795 kWh), Lucknow (182,486 kWh), Shimla (181,498 kWh), New Delhi (176,568 kWh), and Hanumangarh (150,548 kWh). The total yearly GHG emission reduction for all selected six cities of the solar drying system is 279 tCO_2, which might be increased by extending the area of the Solar Collector, and it might be an appealing parameter for achieving CDM project approval. The payback period of Bangalore city is 1.3 years and Hanumangarh has a payback period of 1.9 years. The only limitation of this study is that there is uncertainty between the modeling findings and the actual energy yield because several factors can vary depending on location, such as variations in initial cost at different locations, variations in debt ratio, variations in debt interest rate, variations in debt duration, and variations in GHG emission reduction credit rate. This will have a minor impact on the actual outcomes. However, the average percentage of error between the modeling findings and the actual energy yield RETscreen application was less than 5%, which is considered acceptable for feasibility and simulation investigations. Energy saved may be enhanced by expanding the solar collector's surface area. It is more advantageous to provide subsidies to solar drying system projects to promote the integration of the solar system with industrial food drying systems. Solar drying techniques also provide high-quality dried products that fulfill future food demands. These findings might be extremely beneficial to CDM-approved projects in India.

Acknowledgements The author is highly obliged to the University Grants Commission (UGC) and Ministry of Education, Government of India, New Delhi, for providing the Senior Research Fellowship (SRF).

Data availability My manuscript has no associated data.

Declarations

Conflict of interests The authors declare that they have no conflict of interest.

References

Amer BMA, Gottschalk K, Hossain MA (2018) Integrated hybrid solar drying system and its drying kinetics of chamomile. Renew Energy 121:539–547. https://doi.org/10.1016/j.renene.2018.01.055

Asad M, Mahmood FI, Baffo I et al (2022) The cost benefit analysis of commercial 100 MW solar PV: the Plant Quaid-e-Azam Solar Power Pvt Ltd. Sustainability 14:2895. https://doi.org/10.3390/su14052895

Ashofteh H, Forough AB (2022) Renewable energy's potential scrutiny by PVSYST and RETSCREEN softwares case study: Khoy City, Iran

Aulakh J, Regmi A, Fulton J, Alexander C (2013) Estimating post-harvest food losses: developing a consistent global estimation framework. Agricultural & Applied Economics Association AAEA CAES Jt annual meeting

Castillo Téllez M, Pilatowsky Figueroa I, Castillo Téllez B et al (2018) Solar drying of Stevia (Rebaudiana Bertoni) leaves using direct and indirect technologies. Sol Energy 159:898–907. https://doi.org/10.1016/j.solener.2017.11.031

Ekka JP, Palanisamy M (2021) Performance assessments and techno and enviro-economic analyses on forced convection mixed mode solar dryer. J Food Process Eng 44:1–13. https://doi.org/10.1111/jfpe.13675

El Hage H, Herez A, Ramadan M et al (2018) An investigation on solar drying: a review with economic and environmental assessment. Energy 157:815–829. https://doi.org/10.1016/j.energy.2018.05.197

EL-Shimy M, (2009) Viability analysis of PV power plants in Egypt. Renew Energy 34:2187–2196. https://doi.org/10.1016/j.renene.2009.01.010

Faraj K, Khaled M, Faraj J et al (2020) Phase change material thermal energy storage systems for cooling applications in buildings: a review. Renew Sustain Energy Rev 119:109579. https://doi.org/10.1016/j.rser.2019.109579

Fudholi A, Othman MY, Ruslan MH et al (2011) Techno-economic analysis of solar drying system for seaweed in Malaysia. Recent Res Energy Environ Landsc Archit 15:89–95

Fudholi A, Sopian K, Bakhtyar B et al (2015) Review of solar drying systems with air based solar collectors in Malaysia. Renew Sustain Energy Rev 51:1191–1204. https://doi.org/10.1016/j.rser.2015.07.026

Gustavsson J, Cederberg C, Sonesson U, Emanuelsson A (2013) The methodology of the FAO study: "global food losses and food waste—extent, causes and prevention"—FAO, 2011

Hajar E, Rachid T, Najib BM (2017) Conception of a solar air collector for an indirect solar dryer. Pear Dry Test Energy Procedia 141:29–33. https://doi.org/10.1016/j.egypro.2017.11.114

Hamid ASA, Ibrahim A, Assadeg J et al (2020) Techno-economic analysis of a hybrid solar dryer with a vacuum tube collector for *Hibiscus cannabinus* L fibre. Int J Renew Energy Res 10:1609–1613. https://doi.org/10.20508/ijrer.v10i4.11460.g8045

Hii CL, Rahman RA, Jinap S, Man YBC (2006) Quality of cocoa beans dried using a direct solar dryer at different loadings. J Sci Food Agric 1243:1237–1243. https://doi.org/10.1002/jsfa

Kamarulzaman A, Hasanuzzaman M, Rahim NA (2021) Global advancement of solar drying technologies and its future prospects: a review. Sol Energy 221:559–582. https://doi.org/10.1016/j.solener.2021.04.056

KhaingHnin K, Zhang M, Mujumdar AS, Zhu Y (2019) Emerging food drying technologies with energy-saving characteristics: a review. Dry Technol 37:1465–1480. https://doi.org/10.1080/07373937.2018.1510417

Khatoon S, Kim M-H (2022) Preliminary design and assessment of concentrated solar power plant using supercritical carbon dioxide Brayton cycles. Energy Convers Manag 252:115066. https://doi.org/10.1016/j.enconman.2021.115066

Koçak B, Fernandez AI, Paksoy H (2020) Review on sensible thermal energy storage for industrial solar applications and sustainability aspects. Sol Energy 209:135–169. https://doi.org/10.1016/j.solener.2020.08.081

Kumar A, Das R (2021) Effect of peripheral heat conduction in salt-gradient solar ponds. J Energy Storage 33:102084. https://doi.org/10.1016/j.est.2020.102084

Kumar A, Singh R, Prakash O, Ashutosh (2014) Review on global solar drying status. Agric Eng Int CIGR J 16:161–177

Kushwah A, Kumar A, Gaur MK, Pal A (2021) Garlic dehydration inside heat exchanger-evacuated tube assisted drying system: thermal performance, drying kinetic and color index. J Stored Prod Res 93:101852. https://doi.org/10.1016/j.jspr.2021.101852

Lakhani J, Patel V (2013) Design and development of solar dryer submitted by, pp 0–46. https://doi.org/10.13140/RG.2.2.28719.94880

Lamrani B, Draoui A (2020) Modelling and simulation of a hybrid solar-electrical dryer of wood integrated with latent heat thermal energy storage system. Therm Sci Eng Prog 18:100545. https://doi.org/10.1016/j.tsep.2020.100545

Lamrani B, Draoui A (2021) Thermal performance and economic analysis of an indirect solar dryer of wood integrated with packed-bed thermal energy storage system: a case study of solar thermal applications. Dry Technol 39:1371–1388. https://doi.org/10.1080/07373937.2020.1750025

Lamrani B, Khouya A, Draoui A (2019) Energy and environmental analysis of an indirect hybrid solar dryer of wood using TRNSYS software. Sol Energy 183:132–145. https://doi.org/10.1016/j.solener.2019.03.014

Lee GR, Frearson L, Rodden R (2012a) An assessment of photovoltaic modelling software using real world performance data

Lee K-H, Lee D-W, Baek N-C et al (2012b) Preliminary determination of optimal size for renewable energy resources in buildings using RETScreen. Energy 47:83–96. https://doi.org/10.1016/j.energy.2012.08.040

Lingayat A, Chandramohan VP, Raju VRK (2017) Design, development and performance of indirect type solar dryer for banana drying. Energy Procedia 109:409–416. https://doi.org/10.1016/j.egypro.2017.03.041

Lingayat AB, Chandramohan VP, Raju VRK, Meda V (2020) A review on indirect type solar dryers for agricultural crops—dryer setup, its performance, energy storage and important highlights. Appl Energy 258:114005. https://doi.org/10.1016/j.apenergy.2019.114005

Llavata B, García-Pérez JV, Simal S, Cárcel JA (2020) Innovative pre-treatments to enhance food drying: a current review. Curr Opin Food Sci 35:20–26. https://doi.org/10.1016/j.cofs.2019.12.001

Menon A, Stojceska V, Tassou SA (2020) A systematic review on the recent advances of the energy efficiency improvements in non-conventional food drying technologies. Trends Food Sci Technol 100:67–76. https://doi.org/10.1016/j.tifs.2020.03.014

Moses JA, Norton T, Alagusundaram K, Tiwari BK (2014) Novel drying techniques for the food industry. Food Eng Rev 6:43–55. https://doi.org/10.1007/s12393-014-9078-7

Natural Resources Canada (2005) Clean energy project analysis
Nwakuba N, Okafor VC, Okorafor OO (2020) Techno-economic analysis of a hybrid solar-electric dryer. Energy Sources Part A Recover Util Environ Eff 00:1–25. https://doi.org/10.1080/15567036.2020.1782537
Owolabi AB, Nsafon BEK, Roh JW et al (2019) Validating the techno-economic and environmental sustainability of solar PV technology in Nigeria using RETScreen Experts to assess its viability. Sustain Energy Technol Assessments 36:100542. https://doi.org/10.1016/j.seta.2019.100542
Perera CO, Rahman MS (1997) Heat pump dehumidifier drying of food. Trends Food Sci Technol 8:75–79. https://doi.org/10.1016/S0924-2244(97)01013-3
Pirasteh G, Saidur R, Rahman SMA, Rahim NA (2014) A review on development of solar drying applications. Renew Sustain Energy Rev 31:133–148. https://doi.org/10.1016/j.rser.2013.11.052
Radosevic N, Liu G-J, Tapper N et al (2022) Solar energy modeling and mapping for the sustainable campus at Monash University. Front Sustain Cities. https://doi.org/10.3389/frsc.2021.745197
Rathore NS, Panwar NL (2011) Design and development of energy efficient solar tunnel dryer for industrial drying. Clean Technol Environ Policy 13:125–132. https://doi.org/10.1007/s10098-010-0279-3
RETScreen®, Centre CEDS RETScreen Software ® online user manual
Selvanayaki S, Sampathkumar K (2017) Techno-economic analysis of solar dryers. In: Green energy and technology, pp 463–493
Sharma SD, Iwata T, Kitano H, Sagara K (2005) Thermal performance of a solar cooker based on an evacuated tube solar collector with a PCM storage unit. Sol Energy 78:416–426. https://doi.org/10.1016/j.solener.2004.08.001
Singh S, Anand A, Shukla A, Sharma A (2020) Technical, financial, and environmental feasibility of solar water heater for residential, commercial, and industrial application: a theoretical approach. Mater Sci Energy Technol 3:648–671. https://doi.org/10.1016/j.mset.2020.07.001
Singh S, Anand A, Shukla A, Sharma A (2021) Environmental, technical and financial feasibility study of domestic solar water heating system in India. Sustain Energy Technol Assessments 43:100965. https://doi.org/10.1016/j.seta.2020.100965
Sivakumar M, Mahalingam S, Mohanraj M (2022) Energy, financial and environmental impact analysis of solar thermal heat pump systems using a direct expansion packed bed evaporator-collector. Sol Energy 232:154–168. https://doi.org/10.1016/j.solener.2021.12.059
Srivastava A, Anand A, Shukla A et al (2021) A comprehensive overview on solar grapes drying: modeling, energy, environmental and economic analysis. Sustain Energy Technol Assess 47:101513. https://doi.org/10.1016/j.seta.2021.101513
Tiwari A (2016) A review on solar drying of agricultural produce. J Food Process Technol. https://doi.org/10.4172/2157-7110.10006
Tiwari G, Katiyar VK, Dwivedi V et al (2013) A comparative study of commonly used solar dryers in India. Int J Curr Eng Technol 3:994–999
Verma S, Das R (2020) Revisiting gradient layer heat extraction in solar ponds through a realistic approach. J Sol Energy Eng 142:1–9. https://doi.org/10.1115/1.4046149
Verma S, Das R (2021) Performance analysis of a solar still driven by a packed bed thermal storage tank during off-sunshine period. J Energy Storage 44:103381. https://doi.org/10.1016/j.est.2021.103381
Verma S, Das R (2022) Evaluation of the effectiveness of base insulation on the productivity of a packed bed solar air heater. J Therm Sci Eng Appl 14:1–10. https://doi.org/10.1115/1.4053742
Verma S, Banerjee S, Das R (2022) A fully analytical model of a box solar cooker with sensible thermal storage. Sol Energy 233:531–542. https://doi.org/10.1016/j.solener.2021.12.035
Vijayan S, Arjunan TV, Kumar A (2016) Mathematical modeling and performance analysis of thin layer drying of bitter gourd in sensible storage based indirect solar dryer. Innov Food Sci Emerg Technol 36:59–67. https://doi.org/10.1016/j.ifset.2016.05.014
Vijayan S, Arjunan TV, Kumar A (2020a) Exergo-environmental analysis of an indirect forced convection solar dryer for drying bitter gourd slices. Renew Energy 146:2210–2223. https://doi.org/10.1016/j.renene.2019.08.066
Vijayan S, Arjunan TV, Kumar A, Matheswaran MM (2020b) Experimental and thermal performance investigations on sensible storage based solar air heater. J Energy Storage 31:101620. https://doi.org/10.1016/j.est.2020.101620
Wang W, Li M, Hassanien RHE et al (2018) Thermal performance of indirect forced convection solar dryer and kinetics analysis of mango. Appl Therm Eng 134:310–321. https://doi.org/10.1016/j.applthermaleng.2018.01.115
Windarta J, Handoko S, Sukmadi T et al (2022) Design of a solar power plant system for micro, small, medium enterprise uses PVSyst 7.1 as a renewable energy alternative source in remote areas. IOP Conf Ser Earth Environ Sci 969:012029. https://doi.org/10.1088/1755-1315/969/1/012029

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Ankit Srivastava[1,2] · Abhishek Anand[1] · Amritanshu Shukla[1,3] · Anil Kumar[4] · Atul Sharma[1]

✉ Abhishek Anand
aanand008@gmail.com

✉ Atul Sharma
asharma@rgipt.ac.in

[1] Non-Conventional Energy Laboratory, Rajiv Gandhi Institute of Petroleum Technology, Jais, Amethi, India

[2] Department of Mechanical Engineering, Feroze Gandhi Institute of Engineering and Technology, Raebareli, India

[3] Department of Physics, University of Lucknow, Lucknow, India

[4] Department of Mechanical Engineering, Delhi Technological University, New Delhi, Delhi, India

Performance Evaluation of Four Meta-Heuristic MPPT Algorithms for a DC-DC Converter based PV System under Dynamic Weather Conditions (DWC)

Vinay Kumar
Department of Electrical Engineering
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Rohini, Delhi, India
vk390677@gmail.com
vinaykumar_2k18ee240@dtu.ac.in

Yash Dagar
Department of Electrical Engineering
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Rohini, Delhi, India
yashdagar22899@gmail.com
yashdagar_2k18ee248@dtu.ac.in

Prakash Chittora
Department of Electrical Engineering
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Rohini, Delhi, India
prakashchittora@gmail.com
prakashchittora@dtu.ac.in

2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP) | 978-1-6654-7258-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICICCSP53532.2022.9862414

***Abstract*—Solar Photovoltaic (PV) Systems are one of the most favorable and advantageous power generation frameworks used worldwide to fulfill the global requirements/needs for Renewable & Clean Energy. The output power characteristics generated by the PV Systems are non-linear and are strongly dependent upon the variable environmental parameters such as the Solar Irradiation (W/m^2) and Temperature (ºC). Each PV System has its own (V_{MPP}, I_{MPP}) point defined at a specific Irradiation (W/m^2) and Temperature (ºC) value at which the power is optimal. Thus, in order to ensure the continuous operation of the PV System at the Maximum Power-Point (MPP), an intelligently smart Maximum Power-Point Tracking (MPPT) mechanism is mandatorily implemented along with the Photovoltaic (PV) System. This paper exhibits a comparative study on the behaviors of the four MPPT algorithms namely: Cuckoo Search Algorithm (CSA), Flower Pollination Algorithm (FPA), Grey Wolf Optimization Algorithm (GWOA), and the Particle Swarm Optimization Algorithm (PSOA) under Dynamic Weather/Environmental conditions based on fluctuations in Temperature (ºC) and Solar Irradiation (W/m^2) for a DC-DC Step-Up Converter based Photovoltaic (PV) System that is modeled & implemented under Matlab/Simulink Environment.**

Keywords—Photovoltaic (PV) System, Maximum Power-Point Tracking (MPPT), Dynamic Weather Conditions (DWC), Cuckoo Search Algorithm (CSA), Flower Pollination Algorithm (FPA), Grey Wolf Optimization Algorithm (GWOA), Particle Swarm Optimization Algorithm (PSOA), Matlab/Simulink Software tool.

I. Introduction

In today's technologically sophisticated global world, the demand and acknowledgment for clean and renewable Energy/Power/Electricity are continuously expanding at an ever-increasing exponential pace, owing to the continued hostile & destructive consequences of global warming and climate change. In comparison with classical Non-Renewable energy sources such as Fossil Fuels and Natural Gases, the Solar Radiant Energy is considered and thought to be a very important and favorable energy source among them, because of it being a clean, pollution-free, reliable, sustainable, and infinitely abundant in nature, Solar Energy poses zero threat to the Environment and the existing Ecosystem. The major drawback of this energy source and PV system technology is the dependency and reliance of the PV Output Power on naturally unpredictable & Dynamic Weather/Climatic & Environmental Conditions. In this paper we have focused on a comprehensive analysis and comparative study of the four Meta-Heuristic MPPT algorithms namely: Cuckoo Search Algorithm (CSA) [1], Flower Pollination Algorithm (FPA) [2], Grey Wolf Optimization Algorithm (GWOA) [3], and Particle Swarm Optimization Algorithm (PSOA) [4] for a Vikram Solar Eldora PV Module (60 Cells & 250 W) under the effect of Dynamically abrupt variations in weather parameters such as Temperature (ºC) and Solar Irradiation (W/m^2) for a sample time of 10 seconds. The behavior of each MPPT is intuitively studied under different climatic scenarios: (A) Standard Weather Conditions, (B) Dynamic Weather Conditions: under fixed Temperature and varying Irradiation, and part (C) Dynamic Weather Conditions: under fixed Irradiation and varying Temperature values.

This paper has been organized and presented under the following structure: in which the Section (I): Introduction - Introduces the readers to the goal and setting of the paper, the Section (II) talks about the general working of the PV System, modeling of the single diode PV Cell, Electrical Specifications of the employed Vikram Solar Eldora (VSPB.60.250.03) PV Module, and the DC-DC Boost Converter specifications, further Section (III) briefly discusses the strategies of the four MPPT methodologies, furthermore, Section (IV) presents the Simulation results and Performance Analysis of the PV System under different Dynamic Weather Conditions (DWC) finally the Conclusion in Section (V) followed by Acknowledgement and References respectively.

II. PV System: Model and Characteristics

A typical block design/diagram illustrating the operating mechanism of a Photovoltaic (PV) System is presented in Fig. 1, comprising a Solar PV Panel, a DC-DC Converter, MPPT Control Algorithm, a PWM Control, and a Load. The solar energy illuminating the PV Panel generates PV current (I_{pv}) and PV Voltage (V_{pv}) inside the PV Panel/Module. To transfer the MPP Power from the PV Panel toward the Load, the MPPT control block absorbs the required parameters (I_{pv}, V_{pv}) based on which it in turn generates the command signal to obtain the best optimal Duty Cycle (D) such that Optimal MPP output power is achieved under all types of possible dynamic weather/climatic scenarios (DWCs) discussed here.

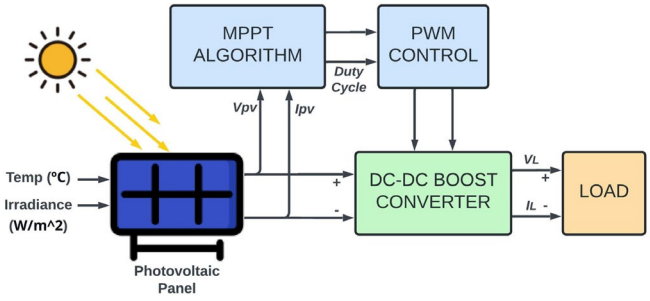


Fig. 1. General Block Design of a PV System.

978-1-6654-7258-6/22/$31.00 ©2022 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 29,2022 at 08:01:17 UTC from IEEE Xplore. Restrictions apply.

The DC-DC Boost/Step-Up Converter employed here acts as an adaptation link to ensure an efficient Maximum Power Transfer from the PV Panel to the connected DC Load. The Boost Converter implements the altered optimum value of the duty cycle adjusted by the MPPT by making the Initial impedance equal to the load impedance. As a result of which the PV System reaches the Maximum Power-Point (MPP).

A. Mathematical Modelling of a Single Diode PV Cell

A PV cell is a semiconductor device, which when exposed to the Sun's Radiant Energy under high-intensity photons falling on the panel surface - creates a huge set of electron-hole pairs in the PN Junction, which further stimulates the movement of electrons and holes across the PN Junction, that in-turn generates the electric current which is known as the Photocurrent (I_{ph}). The Equivalent Circuit Schematic of the widely used single diode PV Cell Model constructed by the interconnection of a Current Source (I_{ph}), Parallel/Shunt Resistance (R_p), Series Resistance (R_s), and a Diode (D) has been depicted in Fig. 2. The electrical elements (I_{ph}, R_p, R_s, D) used in the schematic signifies the Photocurrent, Leakage current, the resistivity of the overall PV material, and electron-hole pair recombination, respectively. The PV cell model of the single diode has been employed here to ensure simplicity and to avoid complexity in project work and its result analysis.

The following mathematical equations validate the design for the one-diode PV Cell model depicted in Fig. 2:

Applying KCL in the PV Cell schematic (Fig. 2), the Output PV Current (I_{pv}) is given by the Eqn. (1)

$$I_{pv} = I_{ph} - I_d - I_{sh} \tag{1}$$

The Current (I_d) flowing through the Diode as leakage current is given by Eqn. (2)

$$I_d = I_o \left(e^{\frac{V_{pv} + I_{pv} R_s}{\alpha V_T}} - 1 \right) \tag{2}$$

The Current Losses (I_{sh}) flowing through the shunt/parallel resistance (R_p) is given by Eqn. (3)

$$I_{sh} = \frac{V_{pv} + I_{pv} R_s}{R_p} \tag{3}$$

The Output PV Characteristic Current (I_{pv}) is given and calculated by Eqn. (4)

$$I_{pv} = I_{ph} - I_o \left(e^{\frac{V_{pv} + I_{pv} R_s}{\alpha V_T}} - 1 \right) - \frac{V_{pv} + I_{pv} R_s}{R_p} \tag{4}$$

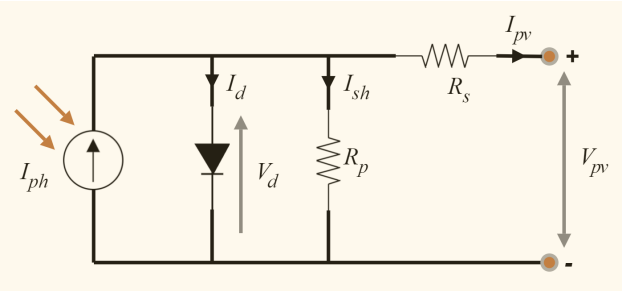


Fig. 2. Circuit Schematic of a One-Diode PV Cell.

Knowingly that,

(I_{pv}) = PV Output Current, (V_{pv}) = PV Open Circuit Voltage, (I_{ph}) = Photocurrent, (I_{sh}) = Shunt Current, (R_p) = Parallel Resistance, (R_s) = Series Resistance, (I_d) = Diode Current, (I_o) = Reverse Saturation Current, (V_T) = Thermal Voltage, q = Electronic Charge (1.67×10^{-19} C), (N_s) = Number of connected PV Cells (Series), K = Boltzmann Constant (1.3805×10^{-23} J/K), T = Temperature (STC).

B. Electrical Parameters of the PV Module/Panel

The Photovoltaic (PV) System used here embodies a Vikram Solar Eldora (VSPB.60.250.03) 250 Watts Photovoltaic module by Vikram Solar Ltd. constituted by 60 PV cells mounted in series and parallel combination with each other. Moreover, the employed PV module in the PV System (Fig. 3) is connected in parallel with the bypass diode which will bypass the PV module in order to avoid the heating effect caused due to the hotspot formation, which is a phenomenon under the effect of Partial shading Irradiation Condition that can damage the PV module permanently. Various electrical characteristics & parameters of the Vikram Solar PV module employed here are summarized in Table I.

TABLE I.
ELECTRICAL PARAMETERS OF THE SOLAR 250 W PV MODULE

Serial No.	Vikram Solar Eldora (VSPB.60.250.03)		
	Parameters/Characteristics	***Abbreviation***	***Ratings***
1.	Maximum Power	P_{max} *(Watts)*	250.14
2.	Voltage at Maximum Power	V_{max} *(Volts)*	30.58
3.	Current at Maximum Power	I_{max} *(Amps.)*	8.18
4.	Open Circuit Voltage	V_{oc} *(Volts)*	37.45
5.	Temperature Coeff. for V_{oc}	α *(%/°C)*	-0.33
6.	Short Circuit Current	I_{sc} *(Amps.)*	8.7
7.	Temperature Coeff. for I_{sc}	β *(%/°C)*	0.0624
8.	Number of Cells per Module	N_{cell} *(Units)*	60

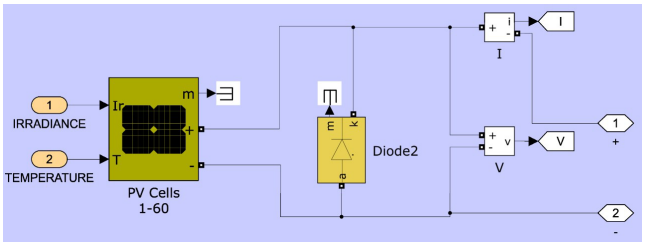


Fig. 3. PV Module (VSPB.60.250.03) Connected to the Bypass Diode.

C. DC-DC Boost (Step-Up) Converter for the PV System

A DC-DC Boost (Step-Up) Converter with a Switching Frequency (F_s) of around 10 kHz, comprising two Capacitors C1 & C2 each of 100 µF capacitance, an Inductor L of 3mH inductance, one switching device (IGBT/MOSFET), and a Diode has been realized as a juncture between the Photovoltaic (PV) System and the DC Load, to boost the incoming voltage & implement the optimally adjusted Duty Cycle (D) by the MPPT control Algorithm, which results in successful MPP transfer from Solar PV Panels to the Load (R_L). The DC-DC Converter circuit is presented in Fig. 4.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 29,2022 at 08:01:17 UTC from IEEE Xplore. Restrictions apply.

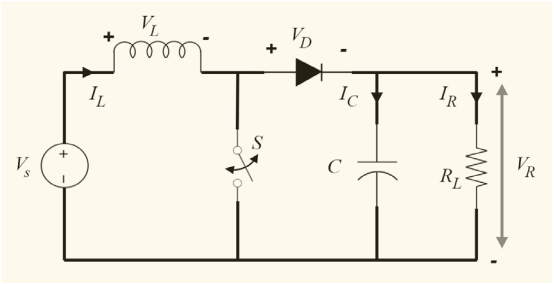


Fig. 4. Equivalent Schematic Circuit of DC-DC Step-UP Converter.

D. I-V & P-V Characteristics of the PV System

Current vs Voltage (I-V) and Power vs Voltage (P-V) characteristics plots for the Vikram Solar Eldora (VSPB.60.250.03) Photovoltaic (PV) Module under Standard Weather Conditions i.e, under fixed Temperature (25ºC) and fixed irradiation (1000 W/m^2) are depicted in Fig. 5. The ideal MPP Power achieved under standard conditions is 250 Watts.

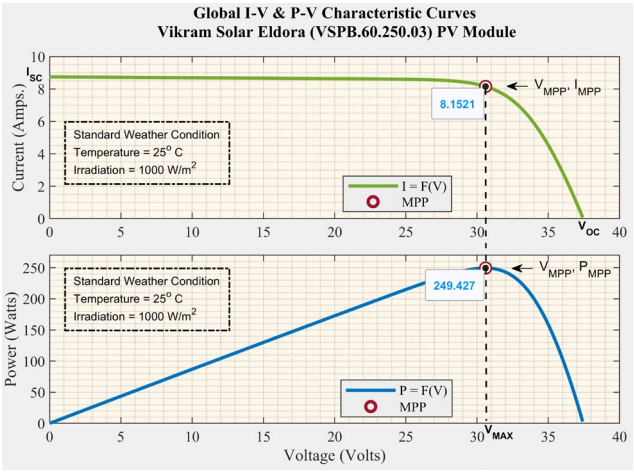


Fig. 5. I-V & P-V Characteristics of the Solar PV Module under STC.

III. The Four Meta-Heuristic MPPT Algorithms

Understanding the extreme importance of a Maximum Power-Point Tracking (MPPT) algorithm in determining and extracting the Global MPP (GMPP) Power from the Photovoltaic (PV) Module in order to make the PV System perform at its best potential, the subsequent sections (A, B, C, and D) present a clear and concise explanation of the working strategy and implementation approach of all the four Meta-Heuristic MPPT techniques/strategies namely CSA, FPA, GWOA, and PSOA to undergo the comparative study of the Vikram Solar PV module/panel under SWC and DWC. [5]

A. Cuckoo Search Algorithm (CSA)

The CSA is a Meta-Heuristic Swarm Intelligence Algorithm originally proposed by Xin-she Yang and Suash Deb in the year 2009 [1], which mimics the nature of Cuckoo bird species, these birds are parasitic in nature as they lay their eggs in the nests built by host bird species in order to let their broods be unwittingly raised by the host parents, which gradually expands the cuckoo generation fertility rate and life expectancy. The odds of the host birds discovering the alien eggs are quite low, as, over time, the cuckoos have evolved the ability to lay eggs that are identical to certain host bird species. The probability of the cuckoo eggs being discovered and tossed away by the host bird is defined as: $P_a \in (0,1)$.

This algorithm is nothing but a simulation of the Levy flight random walk search done by the Cuckoo bird in search of host nests. This algorithmic process begins by initially defining the population for the host nest, and then calculating their respective fitness values. Then, using Levy Flight, new solutions for each host nest are developed, using which the cuckoo keeps on updating its aerial location until the Optimal Global Best point is reached. This approach continuously searches for better solutions and replaces the less optimal ones with better ones to eventually achieve the optimal point.

Cuckoo Search Algorithm is implemented using Eqn. (5)

$$x_i^{t+1} = x_i^t + \alpha \oplus Lévy\ (\lambda) \tag{5}$$

The Levy's distribution (λ) can be calculated by Eqn. (6)

$$Levy \sim u = t^{-\lambda} \tag{6}$$

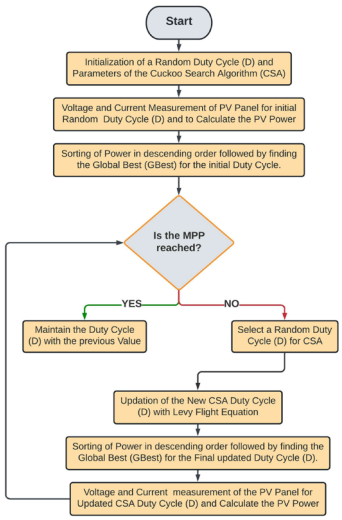


Fig. 6. Flowchart for Cuckoo Search Algorithm (CSA) based MPPT

B. Flower Pollination Algorithm (FPA)

FPA is a Nature-Inspired Meta-Heuristic Algorithm which was originally devised in 2012 by Xin-She Yang. [2] This technique is inspired by the Pollination (Abiotic & Biotic) experience of the blossoming plants. Only roughly 10 % of the blooms perform Self/Abiotic pollination, and the remaining undergo Cross/Biotic Pollination. For Biotic-pollination, numerous pollinators like bees, moths, beetles, bats, insects, etc. are required. The pollinators help the plant in the transfer of the pollens from the anther to the stigma of the flower. Now, to reach the optimal point, a random flower population N is initialized, subsequently, each flower's Fitness is evaluated in the search space, using the fitness function to figure out the best pollen, which is considered the solution. Then, a switch probability is employed to decide between the biotic and abiotic pollination. In the first rule: for global optimum pollination, biotic pollination comes into play, in which the Lévy's Flight (8) is used to carry the pollen by pollinators. The mathematical model for global search in FPA is represented by Eqn. (7). [6]

$$x_i^{t+1} = x_i^t + L(g^* x_i^t) \tag{7}$$

$$\frac{\lambda\Gamma(\lambda)\ sin\ (\pi\lambda/2)}{\pi}\ \frac{1}{s^{1+\lambda}} \tag{8}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 29,2022 at 08:01:17 UTC from IEEE Xplore. Restrictions apply.

Similarly, In the second rule: for Local optimal Pollination, the self-pollination comes into play, we can perform the Local pollination in FPA using Eqn. (9) which mimics the flowering constancy in the limited surroundings.

$$x_i^{t+1} = x_i^t + \varepsilon(x_j^t - x_k^t) \tag{9}$$

Where, x_i^{t+1} = New Solution, x_i^t = Old Solution, g^* = Best Solution, L = Levy's Distribution, x_j^t & x_k^t = Pollens.

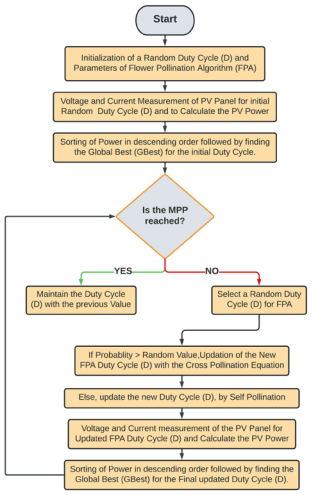


Fig. 7. Flowchart for Flower Pollination Algorithm (FPA) based MPPT

C. *Grey Wolf Optimization Algorithm (GWOA)*

The GWOA is a high-level problem-independent swarm intelligence technique that is inspired by the social hierarchy and hunting techniques used by grey wolves. The GWOA methodology was originally fabricated by Mirjaliali Mohammad and Lewis in the year 2014. [3]

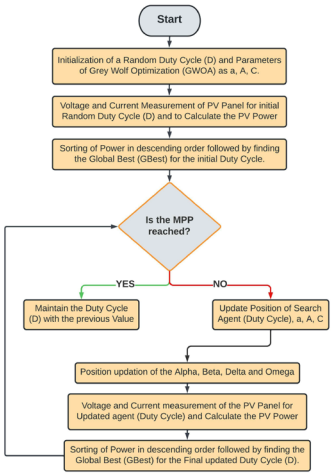


Fig. 8. Flowchart for Grey Wolf Optimization (GWOA) based MPPT

Grey wolves are considered the most successful animals on earth because of their excellent bonding, unity, and loyalty among pack members. Which makes them the survivors of the hardest scenarios. There are generally four ranks in a pack of grey wolves, i.e., Alpha (α), Beta (β), Delta (δ), and Omega (ω). The GWOA mimics and incorporates the leadership and hunting mechanisms of the grey wolves. The approaches used by grey wolves in hunting are (i) Searching, (ii) Tracking & Chasing, (iii) Pursuing & Encircling, and finally (iv) Attacking the Prey. The GWOA's mathematical modeling is validated by Eqn. (10) and (11), which represent grey wolf encircling behavior and that is used to update the grey wolf positions following the prey positions in GWOA.

$$\vec{A} = | \vec{C}.\overrightarrow{X_p} - \vec{X}(t) | \tag{10}$$

$$\vec{X}(t+1) = | \overrightarrow{X_p}(t) - \vec{A}.\vec{D} | \tag{11}$$

Where, t = Current Iteration, $\overrightarrow{X_p}$ = Prey Position, $\vec{X}$ = Grey wolf Position, $\vec{A}, \vec{C}$ = Coefficient vectors.

D. *Particle Swarm Optimization Algorithm (PSOA)*

The PSO Algorithm (PSOA) is a Meta-Heuristic methodology under the category of swarm intelligence (SI). Furthermore, the PSOA technique is a population-based probabilistic time-dependent search algorithm, that is motivated by the social demeanor of a moving flock of birds in a crowd. The PSOA computational strategy was first established by Kennedy and Eberhart in the year 1995. [4] The PSOA methodology works on the population of particles, where every particle is itself a point of a solution. The implementation of this algorithm requires only a few parameters, making it straightforward and simple to solve real-world problems. Starting with a randomly initialized population of candidate particles moving in a randomly chosen direction, each particle goes through the searching space and remembers its best previous position (PBest) and of its swarm neighbors (i.e., Global Best). Each particle searches for the optimum point by continuously updating its generations. In each iteration, every particle finds two best values: (i) PBest and (ii) GBest. The swarm particles update their positional (x_i) and velocity (v_i) values by Eqn. (12) and (13) respectively. PSOA is a simple yet strong algorithm.

New Position of the Particle = Old Position + Velocity

$$x_i^{t+1} = x_i^t + v_i^t * t \tag{12–13}$$

$$v_i^{k+1} = wv_i^k + c_1 r_1 (PBest_i^t - x_i^t) + c_2 r_2 (GBest_i^t - x_i^t)$$

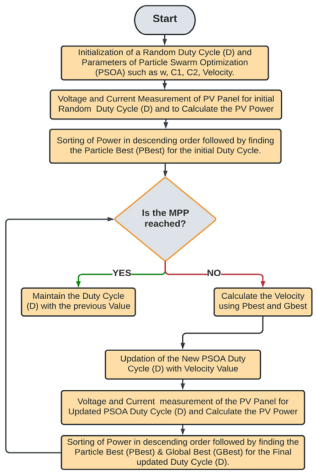


Fig. 9. Flowchart for Particle Swarm Optimization (PSOA) based MPPT

Where, $PBest$ = Particle Best, $GBest$ = Global Best, w = Inertia Weight, c_1, c_2 = Constants, r_1, r_2 = random value.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 29,2022 at 08:01:17 UTC from IEEE Xplore. Restrictions apply.

IV. MATLAB SIMULATIONS: RESULT & ANALYSIS

In this simulation, we have considered three different Weather/Climatic conditions for the Vikram Solar PV Panel/Module to operate under, in order to assess the output efficiency and record the behavior of the four Meta-Heuristic MPPT algorithms loaded to track the MPP power out of the PV System. The three weather conditions simulated here are: (A) Standard Weather Conditions (SWC) having Fixed Temperature = 25º C and Irradiation = 1000 W/m^2 further subdivided into – (i) Without MPPT Control and Case (ii) With MPPT Control, (B) DWC: Fixed Standard Temperature = 25º C & Variable Irradiation, and (C) DWC: Fixed Irradiation = 1000 W/m^2 & Variable Temperature values. The Matlab/Simulink Model for PV Simulation under Standard and Dynamic Weather Conditions is presented in Fig 10. [7]

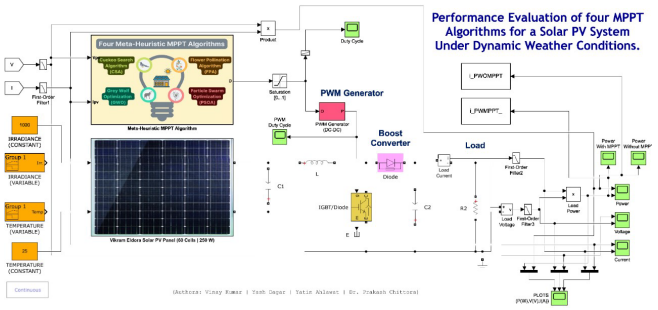


Fig. 10. Matlab/Simulink Model for PV Simulations under SWC & DWC

A. PV System under Standard Weather/Climatic Conditions

The Photovoltaic (PV) System response/behavior under Standard Weather Conditions (SWC) having fixed temperature and fixed illumination (Temp = 25º C & Irradiation = 1000 W/m^2) (i) Without and (ii) With the implementation of MPPT Controls (CSA, FPA, GWOA, PSOA) is depicted in Fig. 11.

It is clearly visible that even under steady-state Standard Weather Conditions (SWC) all four MPPT Algorithms perform differently & better than the PV system operating without MPPT. In order to understand the significance of the MPPT algorithms, MPP Tracking efficiency may be calculated using Eqn. 14 for both (i) & (ii).

$$\eta_{mppt} = \frac{\int_0^t P_{mppt}(t)\,.\,dt}{\int_0^t P_{max}(t)\,.\,dt} \text{ x } 100 \qquad (14)$$

The comparative analogy under STC is not practical but provides us with good insights to have a good understanding of the mechanism and behavior of the MPPT methodologies. In the simulation process, we have considered the simulation time equal to 2 seconds. Resultant MPP Tracking efficiencies (η_{mppt} [%]) calculated by Eqn. 14 for the above (i) without & (ii) with MPPT cases is as follows: Without MPPT (222.67 W, 89.06 %), CSA (228.39 W, 91.35 %), FPA (241.47 W, 96.58 %), GWOA (243.01 W, 97.2 %), PSOA (243.01 W, 97.2 %).

With the use of MPPT, the PV output power has increased by 2.5 %, 8.4 %, 9.1 and 9.1 % respectively. According to the MPP tracking efficiencies, it implies that the MPPT performance is in the order: PSOA > GWOA > FPA > CSA, clearly PSOA outperforms all the other MPPT techniques. This analysis shows the cruciality of the MPPT-based PV systems over PV Systems without MPPT.

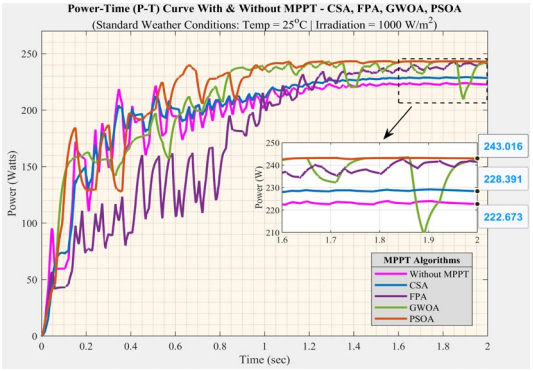
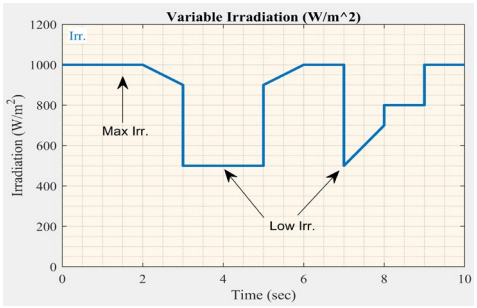


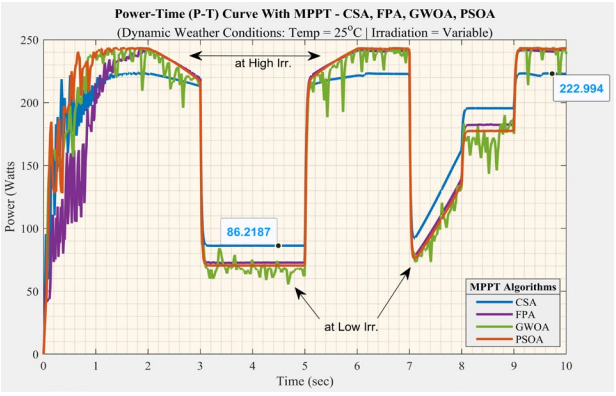
Fig. 11. Optimal MPP Power under Standard Weather Conditions (SWC)

B. PV System under Dynamic Weather Conditions (DWC): Fixed Temperature (ºC) & Variable Irradiation (W/m^2)

In this first case of DWC), we have considered the PV System to be more close to the real/practical scenario, where the Temperature (ºC) is maintained at 25 ºC and solar illumination (W/m^2) to be variable as a linear function of time as depicted in Fig. 12 (i). The Output MPP tracking response of the PV system with MPPT is depicted in Fig. 12 (ii).



(i)



(ii)

Fig. 12. (i) Input and (ii) Output under Fixed Temp. & Variable Irradiance

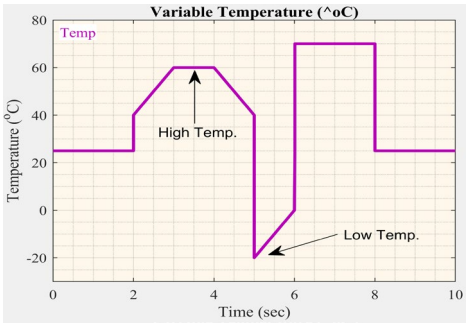
In the foregoing analysis, we have perceived from Fig. 12 (ii) that the P-T curves displayed for four MPPT approaches under the DWC are proportionate to the corresponding dynamic input illumination on the PV panel. Therefore the PV system extracts high MPP power during high irradiation and low MPP power under low irradiation. Moreover, under (a) Lower illumination condition at time 3-5 seconds the Cuckoo Search Algorithm (CSA) outperforms all the other MPPTs but under (b) High Illumination condition at 5-7 seconds the CSA underperforms all the other four algorithms. additionally, the FPA and PSOA techniques show relatively identical performance in extracting the MPP.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 29,2022 at 08:01:17 UTC from IEEE Xplore. Restrictions apply.

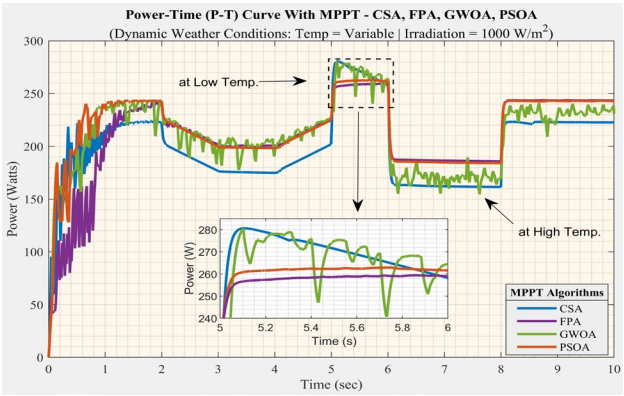
C. *PV System under Dynamic Weather Conditions (DWC): Fixed Irradiation (W/m^2) & Variable Temperature (ºC)*

In the second case of DWC, we have considered the PV System to be under a real-time scenario where the solar Irradiance illumination (W/m^2) is fixed at 1000 W/m^2, and the Temperature (ºC) is varied abruptly as a linear function of time as displayed in Fig. 14 (i). because Temperature is an environmental factor/parameter which varies very slowly in real-time practical settings. The corresponding Dynamic output response of the PV system is depicted in Fig. 13 (ii).

For the PV system to be under the DWC with variable Temperature and fixed Irradiation of 1000 W/m^2, we observe that the P-T curves in Fig. 13 (ii) represent an inversely proportional relationship between the temperature variation and the Output Power, as in Fig. 13 (i) at time 2-5 seconds the Temperature gradually increases and then slowly decreases followed by a sudden fall at the 5th second, corresponding to which, in Fig. 13 (ii) a similar but inverse behavior is seen for the Output MPP Power. Moreover, at the Low temperature of -20 ºC, the CSA has outperformed with 280 W while at a High Temperature of 70ºC the CSA has underperformed in contrast to the FPA, GWOA, and PSOA.



(i)



(ii)

Fig. 13. (i) Input and (ii) Output under Fixed Irradiance. & Variable Temp.

V. Conclusion

In this paper, we have performed a comparative study and analysis on the individual response of the Four Meta-Heuristic MPPT Algorithms, namely CSA, FPA, GWOA, and PSOA for a DC/DC Step-Up Converter based Vikram Solar Eldora (250 W) PV System. In the first part (A) of the simulation, it was observed that the implementation of the MPPT algorithm made the PV system attain up to 9 % of an additional power than what was earlier being achieved under no MPPT scenario. Therefore an MPPT is very crucial for a PV system to perform at its best potential under full capacity. Further, the simulation results achieved in the subsequent parts (B) and (C) under the DWCs were observed and the following insights were drawn: The CSA was found to be efficiently tracking MPP under lower temperatures than in higher temperatures, that too better than the PSOA in the latter low-temperature condition. The GWOA achieved the MPP earlier than even the most efficient MPPT i.e. PSOA, but was unable to acquire the stability in the transient response, rather constantly kept on fluctuating around the optimal point. In terms of the MPP convergence speed of algorithms, the PSOA and FPA were found to be faster than CSA and GWOA. Considering the stability of the transient response - out of the four MPPT algorithms, the PSOA, FPA, and CSA are observed to be stable with minimal fluctuations, while on the other hand GWOA undergoes numerous fluctuations and is unable to attain the final stability. Considering the complexity of their implementation, they have an order based on the number of steps/parameters involved as follows: FPA > GWOA > PSOA > CSA. This comparative analogy of these four Meta-Heuristic MPPT algorithms is meant to aid & help the current researchers in the field of Solar PV energy generation in studying and producing efficient, clean, and long-lasting sustainable solar energy for the future of humankind and life on the earth. [8]

Acknowledgment

The authors of the paper would like to acknowledge and extend their humble & heartfelt thanks with gratitude towards their alma-mater, Delhi Technological University (Formerly DCE), Delhi, and Prof. (Dr.) Prakash Chittora, EED, DTU for their esteemed guidance and constant support in the Project.

References

[1] X. Yang and Suash Deb, "Cuckoo Search via Levy flights," in *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, 2009, pp. 210–214. doi: 10.1109/NABIC.2009.5393690.

[2] X.-S. Yang, "Flower Pollination Algorithm for Global Optimization," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7445 LNCS, 2012, pp. 240–249. doi: 10.1007/978-3-642-32894-7_27.

[3] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[4] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 1995, pp. 39–43. doi: 10.1109/MHS.1995.494215.

[5] M. S. Wasim, M. Amjad, S. Habib, M. A. Abbasi, A. R. Bhatti, and S. M. Muyeen, "A critical review and performance comparisons of swarm-based optimization algorithms in maximum power point tracking of photovoltaic systems under partial shading conditions," *Energy Reports*, vol. 8, pp. 4871–4898, Nov. 2022, doi: 10.1016/j.egyr.2022.03.175.

[6] J. Prasanth Ram and N. Rajasekar, "A Novel Flower Pollination Based Global Maximum Power Point Method for Solar Maximum Power Point Tracking," *IEEE Trans. Power Electron.*, vol. 32, no. 11, pp. 8486–8499, Nov. 2017, doi: 10.1109/TPEL.2016.2645449.

[7] M. Naseem *et al.*, "Assessment of Meta-Heuristic and Classical Methods for GMPPT of PV System," *Trans. Electr. Electron. Mater.*, vol. 22, no. 3, pp. 217–234, Jun. 2021, doi: 10.1007/s42341-021-00306-3.

[8] W. Anjum, A. R. Husain, J. Abdul Aziz, S. M. Fasih ur Rehman, M. P. Bakht, and H. Alqaraghuli, "A Robust Dynamic Control Strategy for Standalone PV System under Variable Load and Environmental Conditions," *Sustainability*, vol. 14, no. 8, p. 4601, Apr. 2022, doi: 10.3390/su14084601.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 29,2022 at 08:01:17 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/361420181

Phycoremediation of Dyes: An Overview on Mechanism, Challenges, and Prospects

Article *in* Ecology, Environment and Conservation · June 2022

DOI: 10.53550/EEC.2022.v28i02.026

CITATIONS
0

READS
20

2 authors, including:

Mohita Chugh
Delhi Technological University
9 PUBLICATIONS **29** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Commercial Hydroponics Set-up for Industrial effluent treatment: Fully funded CST, UP, Govt. of India View project

Cost-effective Hydroponics system: Fully funded by CST, UP, Govt. of India. View project

All content following this page was uploaded by Mohita Chugh on 18 July 2022.

The user has requested enhancement of the downloaded file.

Eco. Env. & Cons. 28 (2) : 2022; pp. (752-757)
Copyright@ EM International
ISSN 0971–765X

DOI No.: http://doi.org/10.53550/EEC.2022.v28i02.026

Phycoremediation of Dyes: An Overview on Mechanism, Challenges, and Prospects

M. Chugh[1] and N. Bharadvaja[2]*

Department of Biotechnology, Delhi Technological University, Delhi 110 042, India

(Received 5 July, 2021; Accepted 7 August, 2021)

ABSTRACT

Dyes are major water pollutants as they show hazardous effects towards humans and aquatic environment. Various physical and chemical methods are used for the remediation of dyes but in the last few years, focus has been shifted towards the modern technique of remediation. Biological techniques have been exploited and utilized for the remediation of dyes as they are more sustainable, simple, and lead to the development of green technology. This review presents the different microalgal strains which are used for the remediation of dyes. It is seen that alone algal bioremediation of dyes is not an efficient and economical process. To overcome this challenge, a more advanced system, phyconanoremediation of dyes is also discussed briefly. Algal nanoparticles can achieve more remediation within few hours which makes this technique more efficient. Moreover, to make this more economical, a strategy to use the left-over biomass to produce valuable products like bioelectricity, pigments, biofuel and other medical products is also introduced. Later, challenges and opportunities are also discussed to make algal bioremediation more effective, sustainable, green, and clean technology.

Keywords *: Algal Bioremediation, Dyes, Phyconanoremediation, Challenges, Prospects.*

Introduction

Industrial effluents, domestic discharge and agrochemical residues are the main sources of water pollution Chaukura *et al.* (2016). Industries like paints, pigments, and textiles release dyes in the water stream. Dyes are aromatic synthetic compounds and have toxic impacts on the ecosystem. Synthetic dyes are used in numerous sectors like in food, leather, cosmetics, electronic industry, pharmaceuticals, printing, plastic, etc. These dyes are seen by the naked eyes in the water bodies but are difficult to abate because of their aromatic structure. More than 3000 dyes are used in the textile, printing and various other purposes and effluent of these industries are released in an uncontrolled manner and pollute the environment (Piaskowski *et al.*, 2018). These have harmful impact on aquatic systems because they decrease the light deep inside the sea. Thus, reducing photosynthetic activity. Many physical, chemical, and biological methods are used for the remediation of dyes like precipitation, sedimentation, adsorption, ion-exchange, bioremediation, etc. These physical and chemical methods are not efficient, eco-friendly, and economical as compared to biological methods as they utilize toxic chemicals and requires energy for the abatement of dyes. To overcome this, scientists are now moving towards the biological methods as they are eco-friendly and more effective as compare to physicochemical techniques (Devasena and Thiruchelvi, 2019). Remediation of dyes with the help of biological

([1]M.Tech Student, [2]Assistant Professor)

agents is known as bioremediation. Bioremediation carried out by algae is known as Phycoremediation. It is observed that phycoremediation is one of the potential techniques of remediation because of their, eco-friendly nature, ease to grow, handle, and maintain. Algae are the hyper accumulators of dyes. Moreover, algae come with an added benefit as they are the rich source of proteins, carbohydrates, lipids etc., which leads to the production of various commercial products like biofuel, medicines, pigments, natural dyes, etc. (Arora *et al.,* 2016). To make this technique more efficient, biological synthesis of nanoparticles with the help of algae is discussed. Recently, the green synthesis of nanoparticles is gaining more attention because of its efficiency, environment-friendly, simple, and expeditious rate of synthesis Kumar *et al.* (2021). This mini-review explores the biological remediation of dyes with the help of algae and discusses the more advanced and economical system of remediation of dyes. Moreover, challenges and opportunities are also briefly mentioned to make this algal-based bioremediation of dyes more economical, efficient, and promising abatement technique.

Phycoremediation of dyes and its mechanism

Algae can abate dyes with the help of three mechanisms that are biosorption, bioaccumulation, and biodegradation. In biosorption, it is observed that dyes bind to functional groups like hydroxyl, amine, thiol, etc., which are present on the surface of algae through electrostatic attraction Zohoorian *et al.* (2020). Algae adsorb chromophores on its cell surface, and it is a passive process and can be carried out by living or non-living biomass. Contrarily, bioaccumulation and biodegradation are active process and requires energy for the remediation of dyes. In bioaccumulation, dyes are accumulated inside the cell with the help of various transport proteins present on the surface of algae. On the other hand, in biodegradation, dye molecules are converted into non-toxic compounds with the help of enzymes or cell metabolites. In bioaccumulation, dyes are consumed by algae as nutrients. These methods are more economical and compatible as compared to conventional methods. The biological methods produce non-hazardous compounds and consume less quantity of water as compared to conventional methods (Chugh *et al.,* 2020).

Algae can decolorize many dyes under optimal growth conditions like pH, temperature, light, and media. Different types of dyes have different decolorization rates. Azo dyes show a high rate of decolorization as compared to phthalocyanine (Abiri *et al.,* 2017). Ethidium bromide is remediated by *Desmodesmus subspicatus, Chlorella vulgaris and Raphidocelis subcapitata.* For complete remediation of ethidium bromide 10^{10} algae/ml is required for *Desmodesmus subspicatus and Chlorella vulgaris*. *Raphidocelis subcapitata required* 10^{11} algae/ml de Almeida *et al.* (2020). *Chlorella vulgaris* are able to remediate ramazol black B, Malachite green, ramazol red RR, ramazol golden yellow RNL and tectilon yellow 2G Alaguprathana and Poonkothai (2017). *Chlorella pyrenoidosa* and *spirulina maxima* dried biomass show photocatalytic reduction of methylene blue. 98.20% of methylene blue is removed by *Chlorella pyrenoidosa* and 94.19% of methylene blue is abated by *Spirulina maxima*Y. A.R. Lebron *et al.* (2019).

Chlorella marina, Isochrysis galbana, Tetarselmis species, Nannochloropsis species, Dunaliella salina and Chlorella species helps in the dye removal from the textile effluent. Malachite green is removed by various strains of microalgae which are *Nostoc species, Turbinaria conoides* and *Chlorella vulgaris. Spirogyra* species are used for the remediation of Acid orange 7, reactive yellow 22, basic blue 3 and basic red 46 dyes. Crystal violet is remediated by *Chactophora elegans*. *Ulothrix species* are used to adsorb methylene blue. Azo dyes are remediated by *Chlorella pyrenoidosa, Chlorella vulgaris, Spirogyra and Oscillatoria tenuis*. More than 30 azo dyes are completely detoxified by *C. pyrenoidosa ,Chlorella vulgaris* and *Oscillateria* sp. Ajaz *et al.* (2020); Hernández-Zamora *et al.* (2015). More recent studies of dye removal by microalgae are mentioned in Table 1.

Phyconanoremediation: More advance phycoremediation technique

It is observed that phycoremediation is not able to achieve complete remediation and consider to be an expensive method of treatment. To overcome this many advanced techniques, like remediation of dyes with the help of phyconanoparticles, can achieve more than 90% of remediation within 5 to 30 minutes and make algae a potential candidate for remediation. These green nano-sized particles aremoreeco-friendly as compared to the chemical synthesis of nanoparticles as they don't use harmful chemicals for the production of these tiny-units. Algae can synthesize nanoparticles with the help of

Table 1. Phycoremediation of dyes

Algae	DYE	Removal Efficiency	References
Chlorella vulgaris	Naphthol Green B	98.5%	Abd Ellatif *et al.*(2020)
	Brazil Wood	99.5%	Abd Ellatif *etal*. 2020)
	Orange G	99.5%	Abd Ellatif *et al*. (2020)
	Methylene Blue	83.04±2.94%	Chin *et al*. (2020)
Anabaena oryzae	Crystalline Violet	97.4%	Ibrahim *et al.* (2021)
Wollea saccata	Malachite Green	93.3%	Picchio *et al.*(2020)
Chara vulgaris L.	Methyl Red	70%	Mahajan and Kaushal (2020)
Enteromorpha intestinalis	Malachite Green	99.63%	Hamouda *et al.*(2020)
Haematococcus pluvialis	Malachite Green	66.5- 95.2%	Liu *et al.* (2018)
Spirogyra sp. and *Oscillatoria sp*	Red and Blue dye.	*Spirogyra* sp.: 78.29% of blue dye and 64.21% red dye. *Oscillatoria* sp.: 76.48% of blue dye and 62.63% of red dye	Brahmbhatt and Jasrai (2016)
Iridaea cordata	Crystal Violet and Methylene Blue	CV: 75% MB: 90%	Escudero *et al.* (2017)
Fucus vesiculosus	Methylene Blue and Eriochrome Black T	MB: 99.28%; EB T: 99.44%	Yuri Abner Rocha Lebron *et al.* (2019)
Ulva fasciata and *Sargassum dentifolium*	Methylene Blue	Individually, *Ulva fasciata* : 97% *Sargassum dentifolium:* 85.6% And in Combination: 99.2%	Moghazy *et al.* (2019)

metal precursors, and then these purified nanoparticles are used for the abatement of dyes which makes this technique eco-friendly and more efficient Iravani *et al.* (2014). Nanoparticles show maximum removal efficiency due to their unique physical, chemical, and magnetic properties. It is reported that as we go down the scale, surface-to-volume ratio increases, and quantum and tunnelling effect is changed. They have numerous adsorption sites because of these special features. These nano-sized particles help in the excellent abatement of dyes within few hours Saravanan *et al.*(2021). Green synthesis of phyconanoparticles takes place with the help of two pathways i.e., intracellular and extracellular synthesis. In intracellular synthesis, reducing agents which are present inside the cell are used to producenanoparticles. On the other hand, in extracellular synthesis, enzymes or other bioactive compounds are released by algal cells are used for the reduction of metal ions into zero-valent nanoparticles Chaudhary *et al.* (2020). There will be no need for reducing orstabilizing agents as algal functional groups will initiate, synthesize, and stabilized the nanoparticle itself. Thus, making it an environment-friendly and cost-effective technique Negi and Singh (2018).

Chlorella vulgaris synthesize silver nanoparticles that can remediate 96.51% methylene blue within three hours Rajkumar *et al.* (2021). Silver phyconanoparticles are also able to remediate methyl red (Husain *et al.,* 2019), methyl orange and congo red Aboelfetoh *et al.* (2017). On the other hand, gold phyconanoparticles are also used in the photocatalytic degradation of dyes. *Sargassum tenerrimum* can synthesize 5-45 nm gold nanoparticles which are used for the abatement of Rhodamine B and Sulforhodamine 101 Ramakrishna *et al.* (2016). To make this phyconanoremediation more economical, one can use the left-over biomass for the production of various commercial by-products like biofuel, pigments, biofertilizer, nutraceuticals, cosmetic products, therapeutic proteins and other medicines Hu *et al.* (2019). Moreover, these nanoparticles can be recycled and re-used for remediation or for other applications which make this process, more efficient, economical as well as a promising approach for the remediation of dyes as shown in Fig. 1.

Challenges and Prospects

Phycoremediation is an outstanding mechanism for the remediation of dyes. Despite various methods, still it is seen that dyes are not completely or quickly remediated by the microalgae. For complete bioremediation, more advanced technologies like genetic engineering, metabolic engineering or protein engineering could play a significant role in it. To make the phycoremediation process more efficient, it isimportant to carry out studies on the gene level and modify the genetic make-up of the microalgae to enhance the efficacy of the process. With the complete genomic studies, it will be possible to make the biomarkers for the contaminated site. CRISPR-Cas 9 tool can be used to make this process more efficient. Toxicological studies of the intermediates produced after transforming the dyes should be done along with the characterization of the strains. However, the cost-efficacy of this process is still an issue faced by sponsors and scientists. This is the major challenge faced by this technology for the commercialization or scale-up of this process. So, more studies should be carried out in the commercialization of this process to make this process more environment friendly and economical at the same time. The strain should be selected not only for maximum remediation of dyes but also able to accumulate maximum by-products like lipids, proteins, pigments, etc. Strain should have lower nutrient requirements and can be able to survive in the robust conditions. More research should be carried out in the field of chemical modification, biomimetic systems, on the effect of tolerance of dyes and to see the influence of dyes on different valuable products. So, that this promising technique will become more cost-effective and more sustainable powerful tool for remediation of dyes. Genetic engineering can be used to enhance the adsorption of dyes over the surface by altering the gene responsible for the expression of protein on the cell surface or enhances the activity of enzyme responsible for dye tolerance and for the accumulation of some by-products. Further studies are required to develop a technology that simultaneously detects as well as abate the dyes from the environment. Still, a lot needs to be understood in the multi-disciplinary field to make this process a successful technique for remediation of dyes on the large scale.

Conclusion

Algal bioremediation of dyes is sustainable, economical, clean, and green technology as compared to traditional processes. Dyes show hazardous effects on humans, flora, and fauna due to their synthetic aromatic nature. These pollutants are released by various industries like textiles, dyeing, printing. Microalgae are the hyperaccumulator of the dyes, CO_2, and are the source of O_2 in the aquatic atmo-

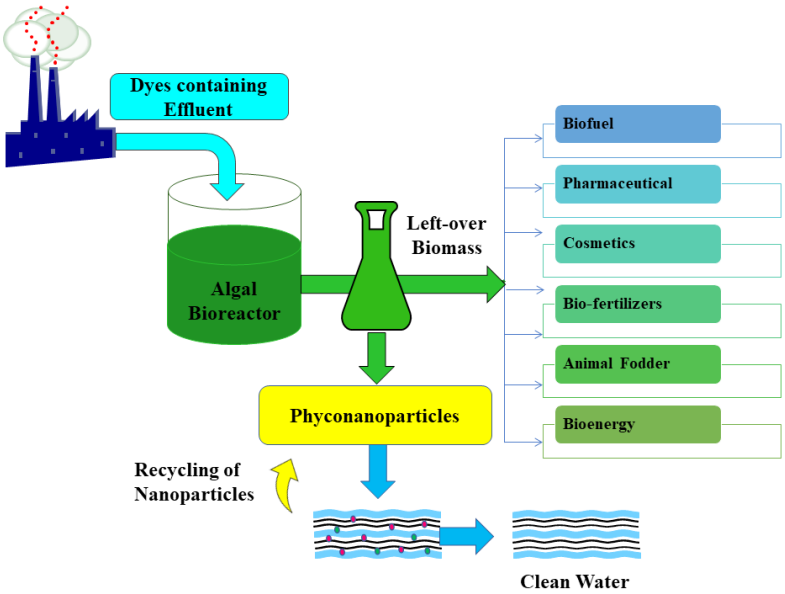


Fig. 1. Economical Phyconanoremediation of Dyes

sphere. Dyes are abated by different mechanisms of action which are biosorption, bioaccumulation, and biodegradation. It is observed that alone phycoremediation is not an effective technique. To overcome, more advanced system, green synthesis of nanoparticles helps in the maximum remediation of dyes within few hours due to their unique properties. These phyconanoparticles are easy to synthesize, maintain and are environment-friendly. Moreover, to make this technique more economical, leftover biomass can be used to produce other economical by-products. Algal species can produce various innovative by-products like biodiesel, biomethane, organic fertilizers and pharmaceutical products. This will make this technique more sustainable, efficient, and economical. Despite all these features, still more studies are required to understand the full abatement mechanism of phyconanoparticles, tolerance of algae towards dyes, effects of dye on the accumulation of different by-products, etc. This will make phycoremediation a powerful tool for the remediation of dyes.

Acknowledgment

The authors would like to thank Delhi Technological University, Delhi, India for providing facilities for the output in this paper form.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Abd Ellatif, S., El-Sheekh, M. M. and Senousy, H. H. 2020. Role of microalgal ligninolytic enzymes in industrial dye decolorization. *International Journal of Phytoremediation*. 0(0) : 1–12. https://doi.org/10.1080/15226514.2020.1789842

Abiri, F., Fallah, N. and Bonakdarpour, B. 2017. Sequential anaerobic-aerobic biological treatment of colored wastewaters: Case study of a textile dyeing factory wastewater. *Water Science and Technology*. 75(6): 1261–1269. https://doi.org/10.2166/wst.2016.531

Aboelfetoh, E. F., El-Shenody, R. A. and Ghobara, M. M. 2017. Eco-friendly synthesis of silver nanoparticles using green algae (Caulerpa serrulata): reaction optimization, catalytic and antibacterial activities. *Environmental Monitoring and Assessment*. 189(7). https://doi.org/10.1007/s10661-017-6033-0

Ajaz, M., Shakeel, S. and Rehman, A. 2020. Microbial use for azo dye degradation—a strategy for dye bioremediation. *International Microbiology*. 23(2) : 149–159. https://doi.org/10.1007/s10123-019-00103-2

Arora, N., Patel, A., Sartaj, K., Pruthi, P. A. and Pruthi, V. 2016. Bioremediation of domestic and industrial wastewaters integrated with enhanced biodiesel production using novel oleaginous microalgae. *Environmental Science and Pollution Research*. 23(20): 20997–21007. https://doi.org/10.1007/s11356-016-7320-y

Brahmbhatt, N. H. and Jasrai, R. T. 2016. The Role of Algae in Bioremediation of Textile Effluent. *International Journal of Engineering Research and General Science*. 4(1) : 443–453.

Chaudhary, R., Nawaz, K., Khan, A. K., Hano, C., Abbasi, B. H. and Anjum, S. 2020. An overview of the algaemediated biosynthesis of nanoparticles and their biomedical applications. *Biomolecules*. 10(11) : 1–36. https://doi.org/10.3390/biom10111498

Chaukura, N., Gwenzi, W., Tavengwa, N. and Manyuchi, M. M. 2016. Biosorbents for the removal of synthetic organics and emerging pollutants: Opportunities and challenges for developing countries. *Environmental Development*. 19 : 84–89. https://doi.org/10.1016/j.envdev.2016.05.002

Chin, J. Y., Chng, L. M., Leong, S. S., Yeap, S. P., Yasin, N. H. M. and Toh, P. Y. 2020. Removal of Synthetic Dye by Chlorella vulgaris Microalgae as Natural Adsorbent. *Arabian Journal for Science and Engineering*. 45(9) : 7385–7395. https://doi.org/10.1007/s13369-020-04557-9

Chugh, M., Kumar, L. and Bharadvaja, N. 2020. Bioremediation of Heavy Metals: A step towards environmental sustainanbility. In *India 2020: Environmental Challenges, policies and green Technology*. (Vol. 66, pp. 77–90).

Devasena, U. and Thiruchelvi, R. 2019. Synthesis of nanoparticles using algae and its application in dye degradation – A review. *International Journal of Research and Analytical Review*. 6(1) : 252–256.

Escudero, L. B., Smichowski, P. N. and Dotto, G. L. 2017. Macroalgae of Iridaea cordata as an efficient biosorbent to remove hazardous cationic dyes from aqueous solutions. *Water Science and Technology*. 76(12) : 3379–3391. https://doi.org/10.2166/wst.2017.505

Hamouda, R. A., El-Naggar, N. E. A., Doleib, N. M. and Saddiq, A. A. 2020. Bioprocessing strategies for cost-effective simultaneous removal of chromium and malachite green by marine alga Enteromorpha intestinalis. *Scientific Reports*. 10(1) : 1–19. https://doi.org/10.1038/s41598-020-70251-3

Hernández-Zamora, M., Cristiani-Urbina, E., Martínez-Jerónimo, F., Perales-Vela, H. V., Ponce-Noyola, T., Montes-Horcasitas, M. del C. and Cañizares-

Villanueva, R. O. 2015. Bioremoval of the azo dye Congo Red by the microalga Chlorella vulgaris. *Environmental Science and Pollution Research*. 22(14) : 10811–10823. https://doi.org/10.1007/s11356-015-4277-1

Hu, J., Guo, H., Xue, Y., Gao, M. tian, Zhang, S., Tsang, Y. F., Li, J., Wang, Y. Nan, and Wang, L. 2019. Using a mixture of microalgae, biochar, and organic manure to increase the capacity of soil to act as carbon sink. *Journal of Soils and Sediments*. 19(11) : 3718–3727. https://doi.org/10.1007/s11368-019-02337-z

Husain, S., Afreen, S., Hemlata, Yasin, D., Afzal, B. and Fatma, T. 2019. Cyanobacteria as a bioreactor for synthesis of silver nanoparticles-an effect of different reaction conditions on the size of nanoparticles and their dye decolorization ability. *Journal of Microbiological Methods*, *162*(December 2018). 77–82. https://doi.org/10.1016/j.mimet.2019.05.011

Ibrahim, M. R., Hamouda, R. A., Tayel, A. A. and Al, M. A. 2021. Anti-cholesterol and Antioxidant Activities of Independent and Combined Microalgae Aqueous Extracts In Vitro. *Waste and Biomass Valorization*. *0123456789*. https://doi.org/10.1007/s12649-021-01349-6

Iravani, S., Korbekandi, H., Mirmohammadi, S. V. and Zolfaghari, B. 2014. Synthesis of silver nanoparticles: Chemical, physical and biological methods. *Research in Pharmaceutical Sciences*. 9(6) : 385–406. /pmc/articles/PMC4326978/

Kumar, L., Ragunathan, V., Chugh, M. and Bharadvaja, N. 2021. Nanomaterials for remediation of contaminants: a review. *Environmental Chemistry Letters*. *0123456789*. https://doi.org/10.1007/s10311-021-01212-z

Lebron, Y. A.R., Moreira, V. R., Santos, L. V. S. and Jacob, R. S. 2018. Remediation of methylene blue from aqueous solution by *Chlorella pyrenoidosa* and *Spirulina maxima* biosorption: Equilibrium, kinetics, thermodynamics and optimization studies. *Journal of Environmental Chemical Engineering*. 6(5): 6680–6690. https://doi.org/10.1016/j.jece.2018.10.025

Lebron, Yuri Abner Rocha, Moreira, V. R. and de Souza Santos, L. V. 2019. Biosorption of methylene blue and eriochrome black T onto the brown macroalgae *Fucus vesiculosus*: equilibrium, kinetics, thermodynamics and optimization. *Environmental Technology (United Kingdom)*. 0(0) : 1–19. https://doi.org/10.1080/09593330.2019.1626914

Liu, J. H., Zhang, L., Zha, D. C., Chen, L. Q., Chen, X. X., and Qi, Z. M. 2018. Biosorption of malachite green onto *Haematococcus pluvialis* observed through synchrotron Fourier-transform infrared microspectroscopy. *Letters in Applied Microbiology*. 67(4) : 348–353. https://doi.org/10.1111/lam.13043

Mahajan, P. and Kaushal, J. 2020. Phytoremediation of azo dye methyl red by macroalgae *Chara vulgaris* L.: kinetic and equilibrium studies. *Environmental Science and Pollution Research*. 27(21) : 26406–26418. https://doi.org/10.1007/s11356-020-08977-w

Moghazy, R. M., Labena, A. and Husien, S. 2019. Eco-friendly complementary biosorption process of methylene blue using micro-sized dried biosorbents of two macro-algal species (*Ulva fasciata* and *Sargassum dentifolium*): Full factorial design, equilibrium, and kinetic studies. *International Journal of Biological Macromolecules*. 134 : 330–343. https://doi.org/10.1016/j.ijbiomac.2019.04.207

Negi, S. and Singh, V. 2018. Algae: A potential source for nanoparticle synthesis. *Journal of Applied and Natural Science*. 10(4) : 1134–1140. https://doi.org/10.31018/jans.v10i4.1878

Piaskowski, K., Œwiderska-D¹browska, R. and Zarzycki, P. K. 2018. Dye removal from water and wastewater using various physical, chemical, and biological processes. *Journal of AOAC International*. 101(5): 1371–1384. https://doi.org/10.5740/jaoacint.18-0051

Picchio, V., Cammisotto, V., Pagano, F., Carnevale, R. and Chimenti, I. 2020. Microalgal Biotechnology: Prospects and Applications. *Intechopen*, *Cell Interaction-Regulation of Immune Responses*, *Disease Development and Management Strategies*. 1–15.

Rajkumar, R., Ezhumalai, G. and Gnanadesigan, M. 2021. A green approach for the synthesis of silver nanoparticles by *Chlorella vulgaris* and its application in photocatalytic dye degradation activity. *Environmental Technology and Innovation*. 21 : 101282. https://doi.org/10.1016/j.eti.2020.101282

Ramakrishna, M., Rajesh Babu, D., Gengan, R. M., Chandra, S. and Nageswara Rao, G. 2016. Green synthesis of gold nanoparticles using marine algae and evaluation of their catalytic activity. *Journal of Nanostructure in Chemistry*. 6(1) : 1–13. https://doi.org/10.1007/s40097-015-0173-y

Saravanan, A., Kumar, P. S., Karishma, S., Vo, D. V. N., Jeevanantham, S., Yaashikaa, P. R. and George, C. S. 2021. A review on biosynthesis of metal nanoparticles and its environmental applications. *Chemosphere*. 264 : 1–16. https://doi.org/10.1016/j.chemosphere.2020.128580

View publication stats

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

Quality of Services in Hybrid SDN (hSDN): A Review

Samiullah Mehraban
Dept. Of Computer Science and Engineering
Delhi Technological University
Delhi, India
mehraban748@gmail.com

R K Yadav
Dept. Of Computer Science and Engineering
Delhi Technological University
Delhi, India
rkyadav@dtu.nic.in

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835774

***Abstract*- As now a day we see that internet has become the backbone of each and every digital society, smart devices as well as the growth of new technology and concepts like IOT, Machine learning 4G, 5G and social network that totally chanced all network requirement. In the last decade Software Define Networking (SDN) changed the Computer Network to an important and significant model as it composed of three different layer Data layer, Control layer and Application layer, main Idea behind SDN is separation of Control layer and Data layer and centralization of control layer to improve efficiency and performance in the network. As Software Define N has lots of benefit compare to legacy network but still deployment of pure SDN is not an easy work and it has its own limitation and some major challenges like financial challenge, technical challenge, security Issue, downtime fear of the network and quality of service degradation. These challenges slow down deployment of fully SDN and cause the birth of Hybrid Software Define Network (hSDN) which legacy devices coexist with SDN devices and take advantage of both legacy and Software Define Network robustness and programmability. In this paper we explore a review of Quality of services in Hybrid SDN in an approach which Quality of Service can benefit the most from the concept of Hybrid Software Define Network.**

Keywords— SDN, hSDN, IOT QOS, 4G ,5G.

I. INTRODUCTION

There has been an unbelievable growth in the telecommunication area in the world in the earlier few years as the development of internet is that much fast that nobody can believe it and has totally changed Information Technology requirements.as we know that the number of mobile and end system is rapidly increasing hence use of internet is rushing via these devices. Modern computer system is packet switched and distributed network system which each and every digital device in the world is connected to each other with the increasing number of mobile and different devices different services has created requirements for network traffic and it has produced needs for flexibility and reliability in the network.

As professional and industries face difficulties in managing todays network and to response this difficulty design of SDN (Software Define Network) started in 2008 as basic Idea behind SDN is separation of Data layer and Control layer. In the past few years Software Define Network is among one of the greatest discussed topics in the telecommunication networks, it is a new and an important concept for network architecture as it appeared in response to the development of Virtualization, mobility, Internet of Things, etc. Software Define Networking is a developed network architecture that separates control layer and data layer[4].

With worldwide view and centralized network control Software Define Network provide reliable and flexible network management, support advance Traffic Engineering (TE), for improvement of link unitization and network throughput of the network it supports smart flow scheduling system. Software Define Network is widely used in different scenario and states like Google [5], public cloud [6], NTT gateway [7], and optical(IP/WDM) network [8]. Software Define Network is used in almost every field of networking from cloud network [9,10] to Data Center [11-13], wide Area Network [14-16] to wireless network [17,18]. According to the survey Software Define Network market is expecting to observe a considerable growth from 8.8 billion USD in 2018 to 28.9 billion USD in 2023, at a Compound Annual Growth Rate (CAGR) 26.8% during this period [19].

Deployment of pure Software Define Network have its own limitation and there are some possible challenges (e.g. financial challenges, technical challenges, lack of standards for migration, security Issue, downtime fears in the network and quality of service degradation) as Software Define Network controllers and switches are not as much stable and reliable and OpenFlow protocol is not developed sufficiently [21]. These challenges hold up deployment of pure Software Define Network and cause the birth of Hybrid Software Define Network (hSDN) which Software Define Network and traditional devices coexist [22].

In Hybrid Software Define Network both traditional devices and Software Define Network devices work and coexist together hence no need to upgrade all of our network devices to Software Define Networks we can upgrade our devices to Software Define Network according to our needs and requirement, hybrid Software Define Network has many benefits and it take advantage of both Software Define Networks and traditional network as it take robustness of traditional network that tradional devices provide robust routing and programmability of Software Define Network as Software Define Network device can emphasis on optimization of the network, in order to provide traffic load balance Software Define Network devices randomly divide traffic flows and divide these flows via multiple path, so network operators has the ability to manage network and balance the load more easily. In the hybrid Software Define Network Controller of network can only focus on traffic or some useful flows, though lots of packets are attained by robustness of traditional network protocols. At the same time, the communication controls that how to control, when to control and for effective and efficient working of the network which devices to connect [23]. different integration level is described in detail in this study [24] like integration in control layer [25], integration in data layer and integration in both data and control layers [26].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

In this Paper we aim to make a picture of Quality of Services in hybrid SDN network, related research studies are organized as motivated in literature according to the categories in a way which quality of service can benefit the most from the concept of hybrid SDN network, scope of this paper is to turn to those terms that QoS characterized by network character like loss, delay, jitter and bandwidth with different industries mechanisms and standards in order to ensure high Quality performance for different application.

We have studied many paper regarding QoS in Hybrid Software Define networking as many research in the literature are done and proposed different methods to achieve , improve and optimize Quality of Service in order to enhance network performance like loss, delay, jitter and link utilization of the network (e.g. minimizing delay, maximizing throughput , balancing link utilization of the network in real time network). but in the very rare case we find that they propose a method to increase all the parameters with good and acceptable percentage as we will focus on maximizing throughput and minimizing delay in Hybrid SDN network.

II. SDN ARCHITECTURE

Software Define Network is an emergent network model that comes with different technologies like Network Function Virtualization and Open Flow protocols. The foremost impression behind Software Define Network architecture intended for implementation and development of agility and cost effectiveness of networks function and network application. Main and key part behind Software Define Network architecture is programmability of the network that we can easily configure and troubleshoot our network within less time, overall Software Define Network is considered by seven important characters that is decoupling of Data layer and control layer, simple devices, central control, automation and virtualization of the network, independence vendor, programmability, and openness and the main or basic idea behind Software Define Network is control layer centralization and decoupling of Data/forwarding layer and control layer for better performance of the network

Fig.1 show the basic architecture of Software Define Network that consist of three main layer that is

Application Layer

Control Layer

Data layer

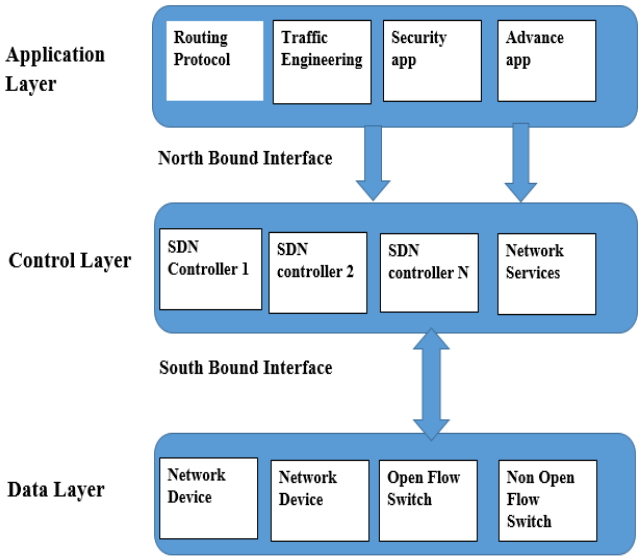


Fig. 1: Software Define Network Basic Architecture

1. Data layer

First or bottom layer is Data layer or infrastructure layer that consist of network devices like routers, access point, physical switches or virtual switches etc. it is responsible for data forwarding also collecting flow statistics and monitoring of network information. Open Flow protocol is the most prevalent standard CDPI (Controller-Data Plane Interfaces) that is used for communication purpose among control layer devices and data layer devices [27-28].

Secure channel is the interface that connect remote controllers to data layer elements. As devices and switches are configured and managed via controller through secure channel.

2. Control layer

Second or central layer of Software Define Network is called control layer which contain one or more than one software based controller which are logically centralized, that configure and manage data plane and carry control functionality of the network. It forward logic table in data layer and control the flow tables. Southbound interface(SBI) is the interface that connect control layer to data layer [29]. Basically controllers contain two important components that is control logic and functional elements as controllers contain more then on function mechanisms like coordinator virtualize etc. controller's logic control map application requirements of networking into orders network element resources [30].

3. Application layer

Third or last layer of Software Define Network is application layer that have one or more than one network application like security application, Routing Protocols, visualization, Traffic Engineering, etc. as it hosts application for introducing new network features, it helps control layer for configuring network with these application requirements. Northbound interface is the interface that used to connect application layer with the control layer [31]. An Open Flow switch in Software Define Network has three different parts: Open Flow Protocol, Flow table and Secure channel

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

III. HYBRID SOFTWARE DEFINE NETWORK

Hybrid Software Define Network is a networking architecture which both legacy devices and Software Define Network devices work together or we can say that centralized and decentralized models coexists and communicate with each other with different degree to control, modify, configure and managing network behavior for optimizing user experience and network performance. as Hybrid Software Define Network take advantage of both Software Define Network and legacy Network, programmability of Software Define Network and robustness of traditional network, no need to upgrade all legacy device to Software Define Network, both Traditional network device and Software Define Network devices work together, as traditional network device provide robust routing, though Software Define Network devices work on optimization of the Network, in traditional network legacy switches control traffic with distributed algorithms like IGP (Interior Gateway Protocol) though in Software Define Network networks Controllers route traffic based on the global view, but in case of Hybrid Software Define Network architecture some part of traffic will be controlled under control of legacy devices and other remaining part of traffic will be controlled by Software Define Network controllers.

In Hybrid SDN network traffics are forwarded based on decision of switching mechanism and distributed routing, control of traffic in Hybrid Software define network depend on various deployment strategies like (Island based, service based, class based, controller based, Hal based, Agent base and overlay based) in some cases traffic control and forwarded through SDN, in some cases some amount of traffic will be controlled through SDN, others controlled with traditional mechanisms, decisions on how to forward and process network traffics are taken by logically centralized controllers.

In Hybrid Software Define Network tradional and legacy devices still use traditional protocols like ECMP (Equal Cost Multipath), Open Shortest Path First (OSPF) in order to forward packets, only Software Define Network device will be controlled by centralized controller. These benefits make Hybrid Software Define Network much attractive to use though hybrid Software Define Network has its own different challenges that is not exists in pure Software Define Network.

According to different analyzed in recent surveys in hybrid SDN network coexistence of different control plane can be source of security issues, as Hybrid Software Define Network is also open to wide range of security issues that inherited from SDN network and traditional network. In real time security id one of the main and important topic in any types of network specially in SDN and hSDN network as it is open and it is centralized network and chance of attack might be increased. Research and surveyed work in this field still need too many work and there are many challenges to be solved it need many more practical research to solve these challenges in order to make Hybrid SDN much better and perfect for every field of networking.

IV. HSDN AND QOS

In today's network supporting end to end quality of services is a big issue and it is an ongoing problem so Researcher have proposed different solution to solve Quality of Service limitation of today's networking models, several of these solutions may failed or either not implemented properly

If we define Quality of service it is a skill of providing required services for a selected network traffic in a network as the key objective of Quality of services is providing precedence with Quality of Service restrictions (bandwidth, loss, delay and jitter) but not limited to these characteristics. To satisfy Quality of Services requirements for different applications SDN provide Open control interface in order to support flexible network traffic scheduling as it is a perfect way to satisfy Quality of Service requirement.

Bi et al. [33] They proposed a scheme that is not only focusing on guarantees of Quality of Services of industrial services but their scheme also focus on an efficient resource bandwidth utilization through traffic load balancing in SDN/OSPF hybrid network. In order to improve QoS for industrial application they propose and intelligent QoS aware forwarding strategy by using SPMC (Single Path Minimum Cost) forwarding scheme and K path partition algorithm for multipath forwarding. For intelligent traffic forwarding they design hybrid network search path (HNSP). In HSNP, they utilize SPMC(Single Oath Minimum Cost) forwarding if the network load is light in OSPF domain, they utilized KPF(K Path Forwarding) algorithm for overloaded links to search multiple path for traffic forwarding. There result shows that their approach guarantees Quality of Services requirements for Hybrid Industrial Internet services, also they efficiently used Open Shortest Path First(OSPF) link bandwidth and balancing traffic load distribution in SDN/OSPF hybrid industrial internet.

Huang et al. [34] for Quality of Services optimization in hybrid Software Define Network they propose near optimal traffic control method. They implemented Flow split ratio approach via locating the Open Flow group bucket restrictions. With open source traffic datasets, they evaluate their proposed method, and their simulation result show that through this method we can achieve an important improvement in network Quality of Services performance. To achieve and improve QoS performance like delay, jitter, loss, and link utilization of network (minimizing delay, balancing link utilization and maximizing throughput) of the network in real time network.

Line et al. [35] proposed QAR(Quality of Services award adaptive routing) in a multilayer hierarchical SDN network. They used distributed hierarchical control plane architecture in large Software Define network for minimizing signaling delay through three level strategy of controllers (Super, domain and slave), with the help of RL(Reinforcement Learning) and Quality of Services provisioning packet forwarding they proposes QAR(Quality of Services aware Adaptive Routing), their result confirm that QAR provides fast convergences and overtakes the existing learning solution with Quality of Services provisioning facilitating practical deployment of SDN in large scale network. QAR algorithm is proposed by the examination of action policy, long term revenue, quality functions and system model with reward functions. Specially to realize efficient adaptive QoS

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

provisioning routing they introduce with MDP (Markov Decision Process) with QoS aware reward function.

Xu et al. [36] To optimize Quality of services for deployment of hybrid Software Define Network under economical constriction they propose heuristic algorithm and proved the approximate factor of l-le. They applied randomized rounding mechanism and depth first search methor for throughput maximization. They show that both simulation algorithms and analysis can obtain an important performance improvement and perform much better than other theoretical cases. there results show that it helps to improve throughput around 40% compared to other scheme and there proposed routing algorithm is able to improve throughput of the network around 31% compared to ECMP (Equal Cost Multipath) in Hybrid SDN network.

Mondel et al. [37] they explore problem of Quality of Services aware data flow management in existence of mixed flows, as this is a Network hard problem they proposed game theory based mixed Data flow management structure, their method has the ability of warranting 24%-47% increase in throughput of the network and can reduce network delay by 77%-98%.

Galan et al. [38] They solve two optimization problem load balance of traffic and network power consumption reduction Their simulation result show which HSLA (Hybrid Spreading Load Algorithm) overtake other methods that challenge one objective only either saving of network power consumption or traffic load balancing.

Guo et al. [39] Their aim was to find best migration system for the traditional routers to migrate to Software Define Network enabled routers that how many routers to migrate to SDN and to decide where to migrate, they propose heuristic algorithm for migration structure for the routers that gets maximum benefit from Traffic Engineering perspective. They evaluate their algorithm by comparing to static migration algorithm and greedy migration algorithm and conducting simulation experiments, there result shows that genetic algorithm is performing very well compared to other techniques in case of properly deployed migration of 40% of router can obtains most of the benefits.

Lin et al. [40] They propose Hybrid Software define network architecture by using Spanning Tree Protocol as they propose Simulated Annealing base Quality of Services aware Routing (SAQR) algorithm that has the ability to change weight of bandwidth and loss rate of delay to find best path for Quality of Services requirements, as they calculate their projected algorithm in a hybrid Software Define Network that runs different application with multiple Quality of Services requirements, Their results show that this algorithm perform much better than other related work MINA in terms of ration of delay, loss and bandwidth with 88%, 90.8% and 86.5% of flows meeting their respective Quality of Services requirements.

Chienhung et al. [41] they propose hybrid Software Define Network architecture that determine presence of traditional network switch using Spanning Tree protocols. They allow open flow Switch to collaborate with traditional switch through using learning bridge protocol without requiring any changes in traditional switches, by using this features of SDN network Software Define Network application can find routing path according to Quality of Service requirements and Current network Status.

Louche et al. [42] proposed geographical routing protocol for vehicular networks with clustering approachs.to send date through selecting best rely it takes three different standard: duration between vehicles, vehicle load, communication error log. This approach allows selection of the most reliable vehicle by ensuring connection availability and avoiding communication problems. When hybrid control layers find next eligible device, data layer is in charge of sending and dividing the data. there model result shows that with respect of packet drops, throughout and average routing overhead, there proposed scheme achieves good performance.

Sanchez et al. [43] They developed MILP (Mixed Integer Linear Programming) optimization Model and their aim was to minimize routing and deployment cost and to maximized controlled traffic in the Software Define network, as they focus on deployments of software define network devices in Hybrid Software Define Network incrementally. There result shows that this method improves the Quality of traffic and rate that managed by Software Define Network model, as they proposed Genetic Algorithm as an execution to achieve comparable performance to the propose MILP (Mixed Integer Linear Programming) model, they showed that Genetic Algorithm is much scalable and strong than Mixed Integer Linear Programming model as traffic volume rises though improving percentage of Software Define Network flows up to 37%.

Siew et al. [44] First in the networking domain they summarize generic work flow for Machine Learning, then from perspective of Machine Learning they study the Problematic implementation of SDN in legacy network and they propose Markov Decision process and Reinforcement Learning Techniques for the deployment problem in Software Define Network. They propose basic workflow of Machine Learning in networking area to formulate Hybrid Software Define Network deployment using Markov Decision process in legacy network, further for model construction and validation Q learning or SARSA technique of reinforcement learning can be used.

Guo et al. [45] They propose RL (Reinforcement Learning) based approach in hybrid Software Define Network to achieve load balancing link they address traffic changing dynamically and learns traffic splitting agent to create direct link between traffic splitting and traffic demands, as traffic splitting agent is designed to learn offline by manipulating Reinforcement Learning, they emulate their work in different topologies and real traffic so there evaluation shows that this method can achieve similar network performance and performs dominance in a fast generation routing schemes.

Table 1 shows literature of the most significant related work Quality of Services and Traffic Engineering (TE) in Hybrid ssoftware Define Networking .

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

Table 1

Traffic Engineering and Quality of Service in Hybrid SDN

Paper	Title /year	Method	Analyze and Result	Conclusion
[33]	Intelligent Quality of Services aware traffic forwarding for SDN/OSPF industrial internet 2020	K Path partition algorithm Single path minimum cost forwarding	They proposed a scheme that is not only focusing on guarantees of Quality of Services of industrial services but their scheme also focus on and efficient resource bandwidth utilization through traffic load balancing in SDN/OSPF industrial internet.	They proposed Quality of Services Aware Forwarding strategy by adopting K path partition algorithm and SPMC (single path minimum cost) algorithm for improvement of industrial application Quality of Services .There result shows that their scheme guarantees Quality of Services requirements for Hybrid Industrial Internet services also they efficiently used Open Shortest Path First link bandwidth and distribution of load balance traffic in SDN/OSPF industrial internet.
[34]	Intelligent traffic control for Quality of Service optimization in hybrid SDN 2021	Deep Reinforcement Learning (DRL)	To optimize Quality of Services in hybrid Software Define Network they propose near optimal traffic control method. They implemented Flow split ratio approach via locating the Open Flow group bucket restrictions.	with open source traffic datasets they evaluate their proposed method and their simulation result show that through this method we can achieve an important improvement in network Quality of Service performance.
[36]	Incremental deployment and throughput maximization routing for hybrid SDN 2017	Heuristic Algorithm, depth first search method	For deployment of hybrid Software Define Networking under economical constriction they propose heuristic algorithm and proved the approximate factor of l-le. They applied randomized rounding mechanism and Depth first search method for throughput maximization.	They show that both simulation algorithms and analysis can obtain an important performance improvement and perform much better than other theoretical cases. there result show that it help to improve throughput around 40% compared to other scheme and there proposed routing algorithm is able to improve throughput of the network around 31% compared to ECMP(Equal cost multipath) in Hybrid Software Define network.
[38]	On the tradeoff B/W load balancing and energy efficiency in hybrid network 2021	Hybrid Spreading Load Algorithm (HSLA)	They solve two optimization problem load balance of traffic and network power consumption reduction .	Their simulation result shows that HSLA(Hybrid Spreading Load Algorithm) overtakes other state of the art methods which challenge one objective only either saving of network power consumption or traffic load balancing
[40]	QoS aware routing in Hybrid SDN network 2017	Simulated annealing based QoS aware routing algorithm	They propose Hybrid Software Define network architecture by using spanning tree protocol as they propose simulated annealing bas Quality of Service aware routing (SAQR) algorithm that has the ability to change weight of bandwidth and loss rate of delay to find best path for Quality of Service requirements, as they calculate their projected algorithm in a hybrid SDN Network that runs different application with multiple Quality of Service requirements	Their results show that this algorithm perform much better rather than other related work MINA in terms of ration of delay, loss, bandwidth with 88%,90.8% and 86.5% of flows meeting their respective Quality of Service requirements.
[43]	Hybrid Incremental Deployment hSDN Devices 2021	Mixed Integer Linear Programming(MILP) Genetic Algorithm (GA)	They developed MILP (Mixed Integer Linear Programming) optimization Model and there aim was to minimize routing and deployment cost and to maximized controlled traffic in the Software Define network, as they focus deployments of Software Define Network devices in Hybrid software define network incrementally.	There result shows that this method improve the Quality of traffic and rate that managed by Software Define network model, as they proposed Genetic Algorithm as an execution to achieve comparable performance to the propose MILP(Mixed Integer Linear Programming) model and they showed that Genetic Algorithm is much more scalable and strong than the MILP(Mixed Integer Linear Programming) model as traffic volume rises though improving percentage of Software Define Network Flows up to 37%.
[45]	Hybrid SDN deployment using Machine learning 2021	RL (Reinforcement Learning)	They propose RL (Reinforcement Learning) based approach in hybrid Software Define network to achieve load balancing link they address traffic changing dynamically and learns traffic splitting agent to create direct link between traffic splitting and traffic demand, as traffic splitting agents designed to learn offline by manipulating Reinforcement Learning.	As they emulate their work in different topologies and real traffic so there evaluation shows that this method can achieve similar network performance and performs dominance in a fast generation routing schemes

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

[44]	Hybrid SDN deployment using ML 2021	Reinforcement learning Techniques	First in the networking domain they summarize generic workflow for Machin Learning , then from the perspective of Machine Learning they study problem of implementing of software define network in tradional network, they proposed Markov Decision process and Reinforcement learning techniques for the deployment problem in software define network.	They propose basic work flow of Machine Learning in networking area to formulate Hybrid Software define network deployment using Markov Decision process in legacy network, further for model construction and validation Q learning or SARSA technique of reinforcement learning can be used.
[39]	Incremental Deployment for traffic engineering in hybrid SDN 2015	Heuristic algorithm	Their aim was to find best migration system for the traditional routers to migrate to Software define network enabled routers that how many routers to migrate to software define network and to decide where to migrate, they propose heuristic algorithm for migration structure for the routers that gets maximum benefit from Traffic Engineering perspective.	they evaluate their algorithm by comparing to Static migration algorithm and greedy migration algorithm and conducting simulation experiments, there result shows that genetic algorithm is performing very well compared to other techniques in case of properly deployed migration of 40% of router can obtains most of the benefits.

In recent years many developing applications like video conferencing, distance learning and many other applications are leading the networking field in the world. Even with the advantage of these Quality of Service needy application for end users it still need work and it still undergo from some issue regarding Quality of Services or Quality of Experience requests to the clients, Quality Of Services between user and service providers mostly relies on Service Level Agreement (SLA) this method work very good for the best efforts services and does not support liner granular traffic control though there are others application which there Flows are sensitive to delay, bandwidth ,jitter like online gaming, VOIP and video conferencing thus requiring special handling technique for these types of applications. Quality of Service are basically applied into two different ways: Hard QoS and Soft QoS. Hard Quality of Service method are used to guarantees the Quality of Service requirement for connection but has resource limitations, Integrated Services (IntServ) is an example of hard Quality of Services, other Quality of Service approach is Soft Quality of Service method that is not strict as much as hard Quality of Service method, differentiated services Diffserv is an example of Soft Quality of Service method.

V. CONCLUSION

The main problem of today network is lack of Quality of Services guarantee as supporting end to end support system is an ongoing problem of today network, Software Define Network come to solve this problem by decoupling of control layer and data layers that provide dynamicity and flexibility to improve network throughout and link utilization of the network and control layer centralization make traffic engineering much easier task. Software Define Network has lots of benefit compare to legacy network but still deployment of pure Software Define Network is not an easy work and it has its own limitation and some challenges like financial challenge, technical challenge, security Issue, downtime fears in the network and quality of service degradation. These challenges slow down deployment of fully SDN and cause the birth of Hybrid Software Define Network (hSDN) which legacy devices coexist with SDN devices and take advantage of both legacy and Software Define Network robustness and programmability. In this review paper we aim to make a picture of Quality of Services in Hybrid SDN network that motivated in literature, we organize related research and surveys studies according to categories in a way which Quality of Service can benefit the most from the concept of Hybrid Software Define Network.

REFERENCES

[1] Khorsandroo, Sajad, et al. "Hybrid SDN evolution: A comprehensive survey of the state-of-the-art." Computer Networks 192 (2021): 107981.

[2] Sinha, Yash, and K. Haribabu. "A survey: Hybrid sdn." Journal of Network and Computer Applications 100 (2017): 35-55.

[3] Huang, Xinli, et al. "A survey of deployment solutions and optimization strategies for hybrid SDN networks." IEEE Communications Surveys & Tutorials 21.2 (2018): 1483-1507.

[4] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, "A survey on software-defined networking," IEEE Communications Surveys & Tutorials, vol. 17, no. 1, pp. 27–51, 2015.

[5] S. Jain, A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu, et al., "B4: Experience with a globally-deployed software defined WAN," ACM SIGCOMM Computer Communication Review, vol. 43, no. 4, pp. 3–14, 2013.

[6] C.-Y. Hong, S. Kandula, R. Mahajan, M. Zhang, V. Gill, M. Nanduri, and R. Wattenhofer, "Achieving high utilization with software-driven WAN," ACM SIGCOMM Computer Communication Review, vol. 43, no. 4, pp. 15–26, 2013.

[7] S. Natarajan, A. Ramaiah, and M. Mathen, "A software defined cloud-gateway automation system using OpenFlow," in Proc. IEEE International Conference on Cloud Networking (CloudNet), pp. 219– 226, IEEE, 2013.

[8] A. S. Thyagaturu, A. Mercian, M. P. McGarry, M. Reisslein, and W. Kellerer, "Software defined optical networks (SDONs): A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 18, no. 4, pp. 2738–2786, 2016.

[9] Mohammad Banikazemi, David Olshefski, Anees Shaikh, John Tracey, Guohui Wang, Meridian: an SDN platform for cloud network services, IEEE Commun. Mag. 51 (2) (2013) 120–127.

[10] Joe Mambretti, Jim Chen, Fei Yeh, Next generation clouds, the chameleon cloud testbed, and software defined networking (sdn), in: 2015 International Conference on Cloud Computing Research and Innovation (ICCCRI), IEEE, 2015, pp. 73–79.

[11] Tao Wang, Fangming Liu, Hong Xu, An efficient online algorithm for dynamic SDN controller assignment in data center networks, IEEE/ACM Trans. Netw. 25 (5) (2017) 2788–2801.

[12] L Velasco, A Asensio, JLl Berral, A Castro, Víctor López, Towards a carrier SDN: An example for elastic inter-datacenter connectivity, Opt. Express 22 (1) (2014) 55–61.

[13] Jing Liu, Jie Li, Guochu Shou, Yihong Hu, Zhigang Guo, Wei Dai, SDN based load balancing mechanism for elephant flow in data center networks, in: 2014 International Symposium on Wireless Personal Multimedia Communications (WPMC), IEEE, 2014, pp. 486–490.

[14] Sushant Jain, Alok Kumar, Subhasree Mandal, Joon Ong, Leon Poutievski, Arjun Singh, Subbaiah Venkata, Jim Wanderer, Junlan Zhou, Min Zhu, Jon Zolla, Urs Hölzle, Stephen Stuart, Amin Vahdat, B4: Experience with a globallydeployed software defined wan, in: Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM, SIGCOMM'13,2013,pp.3–14,.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

[15] Xin Jin, Yiran Li, Da Wei, Siming Li, Jie Gao, Lei Xu, Guangzhi Li, Wei Xu, Jennifer Rexford, Optimizing bulk transfers with software-defined optical WAN, in: Proceedings of the 2016 ACM SIGCOMM Conference, 2016, pp. 87–100.

[16] Chi-Yao Hong, Subhasree Mandal, Mohammad Al-Fares, Min Zhu, Richard Alimi, Kondapa Naidu B., Chandan Bhagat, Sourabh Jain, Jay Kaimal, Shiyu Liang, Kirill Mendelev, Steve Padgett, Faro Rabe, Saikat Ray, Malveeka Tewari, Matt Tierney, Monika Zahn, Jonathan Zolla, Joon Ong, Amin Vahdat, B4 and after: Managing hierarchy, partitioning, and asymmetry for availability and scale in google's software-defined WAN, in: Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication, SIGCOMM '18, 2018, pp. 74–87,

[17] Salvatore Costanzo, Laura Galluccio, Giacomo Morabito, Sergio Palazzo, Software defined wireless networks: Unbridling SDNs, in: 2012 European Workshop on Software Defined Networking, IEEE, 2012, pp.1-6.

[18] Roberto Riggio, Mahesh K Marina, Julius Schulz-Zander, Slawomir Kuklinski, Tinku Rasheed, Programming abstractions for software-defined wireless networks, IEEE Trans. Netw. Serv. Manag. 12 (2) (2015) 146–162.

[19] MARKETSANDMARKETS, Software defined networking market by SDN type (open SDN, SDN via overlay, and SDN via API), component (solutions and services), end user (data centers, service providers, and enterprises), and region - global forecast to 2023, 2019,

[20] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "OpenFlow: enabling innovation in campus networks," ACM SIGCOMM Computer Communication Review, vol. 38, no. 2, pp. 69–74, 2008.

[21] S. Vissicchio, L. Vanbever, and O. Bonaventure, "Opportunities and research challenges of hybrid software defined networks," ACM SIGCOMM Computer Communication Review, vol. 44, no. 2, pp. 70–75, 2014.

[22] L Velasco, A Asensio, JLl Berral, A Castro, Víctor López, Towards a carrier SDN: An example for elastic inter-datacenter connectivity, Opt. Express 22 (1) (2014) 55–61.

[23] Stefano Vissicchio, Laurent Vanbever, and Olivier Bonaventure. Opportunities and research challenges of hybrid software defined networks. ACM SIGCOMM Computer Communication Review, 44(2):70–75, 2014.

[24] Yash Sinha, K Haribabu, et al. A survey: hybrid sdn. Journal of Network and Computer Applications, 100:35 55, 2017.

[25] Matthew Caesar, Donald Caldwell, Nick Feamster, Jennifer Rexford, Aman Shaikh, and Jacobus van der Merwe. Design and implementation of a routing control platform. In Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation-Volume 2, pages 15–28. USENIX Association, 2005.

[26] Cheng Jin, Cristian Lumezanu, Qiang Xu, Zhi-Li Zhang, and Guofei Jiang. Telekinesis: Controlling legacy switch routing with openflow in hybrid networks. In Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research, pages 1–7, 2015.

[27] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, J. Turner, Openow: Enabling innovation in campus networks, SIGCOMM Comput. Commun. Rev. 38 (2) (2008) 69-74.

[28] S. J. Vaughan-Nichols, Openow: The next generation of the network?, IEEE Computer 44 (8) (2011) 13{15.

[29] W. Xia, Y. Wen, C.H. Foh, D. Niyato, H. Xie, A survey on software-defined networking, IEEE Commun. Surv. Tutor. 17 (1) (2015) 27–51.

[30] SDN architecture, Tech. rep., Open Networking Foundation (ONF) (June 2014).

[31] D. Kreutz, F. Ramos, P. Verissimo, C. Rothenberg, S. Azodolmolky, S. Uhlig, Software-defined networking: A comprehensive survey, Proc. IEEE 103 (1) (2015) 14–76.

[32] Mehraban, Samiullah, Komil B. Vora, and Darshan Upadhyay. "Deploy multi protocol label switching (mpls) using virtual routing and forwarding (vrf)." 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI). IEEE, 2018.

[33]Bi, Yuanguo, et al. "Intelligent quality of service aware traffic forwarding for software-defined networking/open shortest path first hybrid industrial internet." IEEE Transactions on Industrial Informatics 16.2 (2019)

[34] Huang, Xiaohong, Man Zeng, and Kun Xie. "Intelligent traffic control for QoS optimization in hybrid SDNs." Computer Networks 189 (2021)

[35] S. Lin, I.F. Akyildiz, P. Wang, M. Luo, QoS-aware adaptive routing in multi-layer software defined networks: a reinforcement learning approach, in: 2016 IEEE International Conference on Services Computing (SCC), 2016, pp. 25–33.

[36] Xu, Hongli, et al. "Incremental deployment and throughput maximization routing for a hybrid SDN." IEEE/ACM Transactions on Networking 25.3 (2017): 1861-1875.

[37] A. Mondal, S. Misra, FlowMan: QoS-aware dynamic data flow management in software-defined networks, IEEE J. Sel. Areas Commun. 38 (7) (2020) 1366–1373.

[38] Galàn-Jimènez, Jaime, et al. "On the tradeoff between load balancing and energy-efficiency in hybrid IP/SDN networks." 2021 12th International Conference on Network of the Future (NoF). IEEE, 2021.

[39] Guo, Yingya, et al. "Incremental deployment for traffic engineering in hybrid SDN network." 2015 IEEE 34th international performance computing and communications conference (IPCCC). IEEE, 2015.

[40] Lin, Chienhung, Kuochen Wang, and Guocin Deng. "A QoS-aware routing in SDN hybrid networks." Procedia Computer Science 110 (2017): 242-249.

[41] Chienhung Lin, Kuochen Wang, Guocin Deng, A QoS-aware routing in SDN hybrid networks, Procedia Comput. Sci. 110 (2017) 242–249, 14th International Conference on Mobile Systems and Pervasive Computing (MobiSPC 2017) / 12th International Conference on Future Networks and Communications (FNC 2017) / Affiliated Workshops.

[42] Lylia Alouache, Nga Nguyen, Makhlouf Aliouat, Rachid Chelouah, HSDN-GRA: A hybrid software-defined networking-based geographic routing protocol with multi-agent approach, Int. J. Commun. Syst. 33 (15) (2020) e4521.

[43]Cespedes-Sanchez, Pedro, et al. "Hybrid Incremental Deployment of HSDN Devices." IEEE Transactions on Network and Service Management (2021).

[44] Siew, Hong Wei, Saw Chin Tan, and Ching Kwang Lee. "Hybrid SDN Deployment Using Machine Learning." Computational Science and Technology. Springer, Singapore, 2021. 215-225.

[45] Guo, Yingya, et al. "Traffic Engineering in Hybrid Software Defined Network via Reinforcement Learning." Journal of Network and Computer Applications 189 (2021): 103116.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:09:36 UTC from IEEE Xplore. Restrictions apply.

Journal of Electronic Materials
https://doi.org/10.1007/s11664-022-09795-1

ORIGINAL RESEARCH ARTICLE

Radiation-Sensitive AlGaN/GaN MOS-HEMT-Based Dosimeter

Ruby Mann[1] · Sonam Rewari[2] · Praveen Pal[3] · Shobha Sharma[4] · R. S. Gupta[5]

Received: 20 March 2022 / Accepted: 22 June 2022
© The Minerals, Metals & Materials Society 2022

Abstract
An AlGaN/GaN MOS-HEMT (metal oxide semiconductor–high electron mobility transistor)-based dosimeter has been proposed to demonstrate and evaluate the impact of absorbed doses of radiations by virtue of interface trap charges at the interface of SiO_2/AlGaN. The effect of various radiation doses induced ($N_t = -3 \times 10^{12}\text{cm}^{-2}$ to $+3 \times 10^{12}\text{cm}^{-2}$) has been studied in terms of various electrical parameters, such as channel conductance (g_d), drain current (I_{ds}), transconductance (g_m), and threshold voltage(V_{th}). Variations in the drain current and channel potential, and a significant shift in the threshold voltage, have been observed. Threshold voltage and drain current increase proportionally to the radiation dose induced due to the addition of positive charges on the surface, leading to increased charge carrier concentration in two-dimensional electron gas (2DEG). Modifying the electron mobility/or density by the effects of induced radiation affects the functionality of the 2DEG-based device. The maximum changes in output conductance, drain current, threshold voltage change, drain-on sensitivity, and transconductance observed for the proposed sensor are 69%, 0.073 A, 593 mV, 0.314, and 0.018 S, respectively. The effects of variation in gate length on the threshold voltage sensitivity and sensor drain current have been studied.

Keywords MOS-HEMT · AlGaN barrier · threshold voltage dosimeter · sensor

Introduction

Demand for FET (field-effect transistors, three terminal devices which use the electric field to control the flow of current through the device)-based sensors has increased in the last few decades. GaN (gallium nitride) is a wide band gap semiconductor material) showing good power density and reliability, and superior overall performance.[1] GaN has been demonstrated to be a suitable material when operated in high-temperature environments and offers stability in the devices.[2] Since GaN was developed, HEMT (high electron mobility transistor) is a field-effect transistor incorporating a junction between two materials with different band gaps as the channel instead of a doped region. Such technology has been used inside mobile base stations in power amplifiers.[3] Due to the wide band gap of 3.43 eV, it affords unique properties for applications in FET sensors.

Moreover, their use has been extended to various radar sensor applications. Under a harsh environment, intense heat,and acidic or basic solutions, the AlGaN/GaN HEMT gives a better performance when used as a sensor.[4] In protein detection, chemicals are appended to the AlGaN/GaN HEMT surface, due to which the current of the device is influenced by the charges attached.[5] AlGaN/GaN HEMT-based sensors deal with the lower theoretical limits and display significant deviations in output current, which is its main benefit.[6] In the present decade, sensors have gained significance in medical and environmental monitoring in less broadly examined applications, for detecting gases, ions, and chemicals, of sensors capable of operating in harsh conditions. Hence, GaN-HEMT sensors are considered better than other FET-based sensors.[7,8] GaN-based FET devices attract

✉ Sonam Rewari
rewarisonam@gmail.com

1 Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, New Delhi 110006, India

2 Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi 110042, India

3 Modeling and Simulation Research Laboratory, Shaheed Rajguru College of Applied Sciences for Women, University of Delhi, New Delhi 110096, India

4 Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, New Delhi 110006, India

5 Department of Electronics and Communication Engineering, Maharaja Agrasen Institute of Technology, New Delhi 110086, India

Published online: 26 July 2022

much attention because they are exceptionally tolerant to high-energy ionizing radiation. Field-effect-based dosimeters such as the metal oxide semiconductor (MOS) dosimeter is of specific interest for various detection applications, such as radiation detection sensors.[9–11] AlGaN/GaN HEMTs possess a two-dimensional electron gas (2DEG) close (15–30 nm) to the surface, which is extremely sensitive to adsorbed gas atoms, molecules, and radiation. Therefore, the GaN-HEMT has stood out as promising for dosimetry application, because of its excellent radiation resistance property.[12,13] The effect of ionizing radiation can be seen clearly on Si microelectronic devices, as it hampers the performance by oxide-trapped charges, bringing out interface traps.[14] Radiation from a source plays a significant role in various applications; for example, radiation therapy used in medicines. So it is vital to calculate the amount of radiation produced by a source, and the amount of radiation absorbed by a substance. When the ionizing radiation penetrates the oxide, it creates a charge that is trapped indefinitely, which causes a change in the threshold voltage.[15–17] MOSFET dosimeters operate on the threshold voltage principle, which is a linear function of the absorbed dose but with a lower sensitivity.[18] Due to the highly sensitive surface of the AlGaN barrier layer, AlGaN/GaN MOS HEMT has been used to develop a dosimeter sensor to evaluate absorbed doses by virtue of the radiation dose induced on the interface of SiO_2/AlGaN.

This paper is organized as follows. In section "Device Fabrication and Simulation," the fabrication and simulation of the device are carried out. In section "Results and Discussion," an investigation of the proposed AlGaN/GaN MOS-HEMT-based dosimeter has been performed. The conclusions are drawn in section "Conclusions."

Device Fabrication and Simulation

Figure 1a shows a schematic view of the AlGaN/GaN MOS-HEMT-based dosimeter. The structure was grown on a sapphire substrate and consisted of a 25-nm AlGaN barrier layer with an aluminium composition of 0.30. In this structure, a 20-nm AIN (aluminium nitride is a non-oxide technical ceramic) nucleation layer has been considered below the GaN buffer layer of 2 µm.The nucleation layer helps to reduce lattice mismatch between the GaN and the substrate layer.[19] The gate width and length are 100 µm and 1 µm, respectively. A SiO_2 layer of 5 nm above the AlGaN barrier layer is considered. The dosimeter is designed to detect radiation in terms of the uniform radiation dose induced.[20] Table I shows all the device parameters with their values used in this paper. A metal–organic chemical vapour deposition system was used to grow the AlGaN/GaN heterostructure on a sapphire substrate. A 20-nm AlN nucleation layer, a 2-µm GaN buffer layer, a 25-nm AlGaN barrier layer, and a 5-nm SiO_2 oxide layer comprised the epitaxial layers. By using Hall effect measurements, the sheet carrier density and 2DEG mobility were determined to be 1.01×10^{12} cm^{-2} and

Table I Device structure parameters

Parameters	Values
AlGaN layer thickness	25 nm
GaN layer thickness	2 µm
Oxide layer thickness	5 nm
Al-mole fraction	0.30
Gate length (L_g)	1 µm
AlGaN/GaN nucleation layer	20 nm

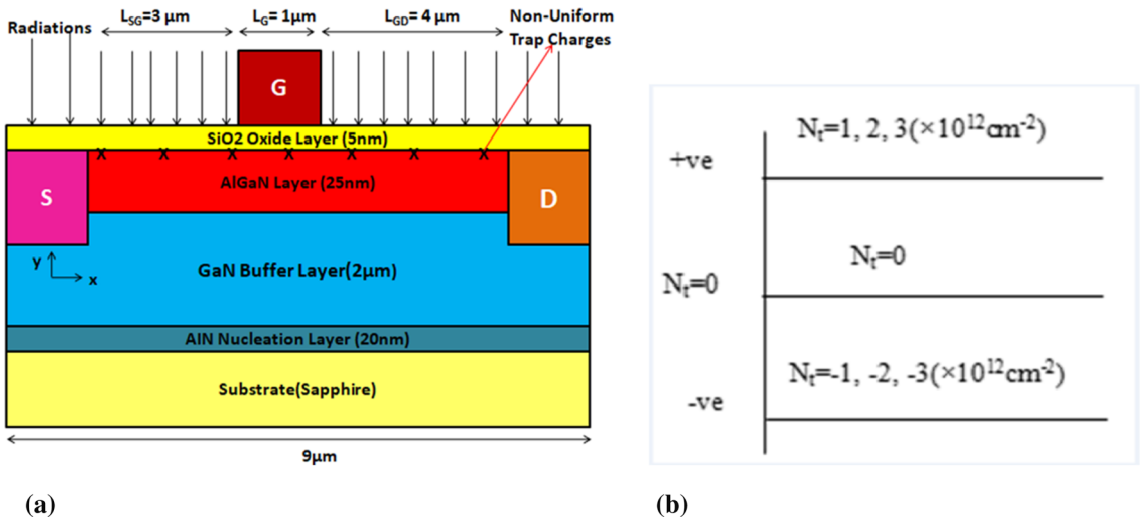


Fig. 1 (a) 2-D structure of AlGaN/GaNMOS-HEMT dosimeter. (b).Uniform charge distribution profile along the channel.

1915 cm^2/V-s, respectively. The source and drain electrodes were created by evaporating Ti/Al/Ni/Au and then annealing for 30 s in a N_2 ambient at 830°C. Following that, mesa isolation was accomplished using Cl-based reactive ion etching. Finally, a Schottky gate was created by evaporating Ni/Au/Ni. The gate-source spacing is 3 µm, and the gate-drain spacing is 4 µm.[21]

Length from source to the gate electrode (L_{SG}) and gate to drain electrode (L_{GD}) are 3 µm and 3.5 µm, respectively. Device performance can be enhanced due to the reduced gate–source distance. A polarization model, field-dependent mobility model, concentration-dependent mobility, and Shockley–Read–Hall models have been used in the simulation. The uniform charge distribution profile along the channel is shown in Fig. 1b.The flowchart of fabrication is given in Fig. 2. Figure 3 displays the $I_{ds}-V_{ds}$ characteristics of the device, with the experimental work carried out and previously reported by Hu et al.[22] It can be seen that the simulated characteristics of $I_{ds}-V_{ds}$ are in good agreement with the experimental data.[22]

Results and Discussion

Potential distribution along the channel for uniform radiation doses induced has been plotted in Fig. 4, from which it can be seen that the highest potential of 1.6062 V is obtained for a uniform radiation dose of $+3\times10^{12}cm^{-2}$. In contrast, a minimum channel potential of 1.6058 V is observed for a radiation dose of $-3\times10^{12}cm^{-2}$. Also, a high change

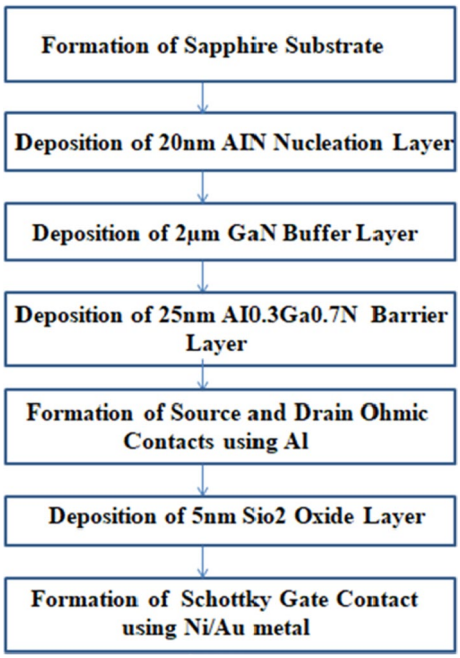


Fig. 2 Flowchart of fabrication.

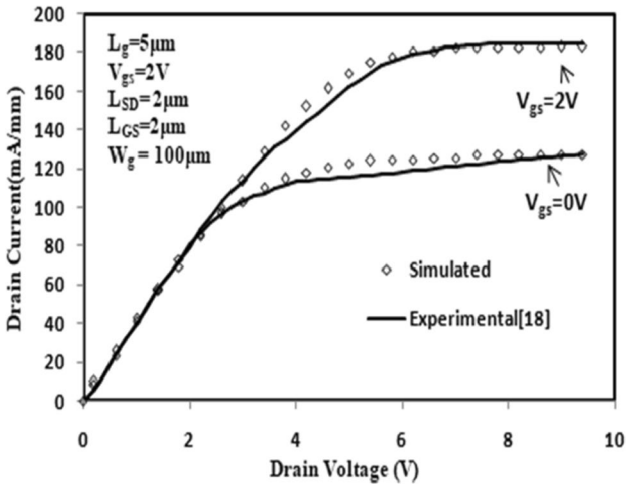


Fig. 3 $I_{ds}-V_{ds}$ characteristics of the MOS-HEMT device.

in potentials can be observed for radiation falling on the source–gate and gate–drain regions. Although, as expected, the metal gate electrode prevents the radiation, a slight change in potential is obtained at the SiO_2/AlGaN interface exactly below the gate. Inducing the interface charges causes changes in the carrier concentration in 2DEG, and leads to change in the electric field at the interface. Considering the negative charge entrapment of the AlGaN layer will decrease the electric field at the interface, allowing a lower electronic density. Hoewver, if the accumulated charge at the interface is positive, this will increase the carrier density and electric field.[23] Figure 5 shows the plot of the electric field along the channel (*x*-axis) of the AlGaN/GaN MOS-HEMT.A high trap charge concentration indicates a high electric field due to increased potential in the 2DEG channel. Maximum and minimum electric field values of 5.15*E*+04

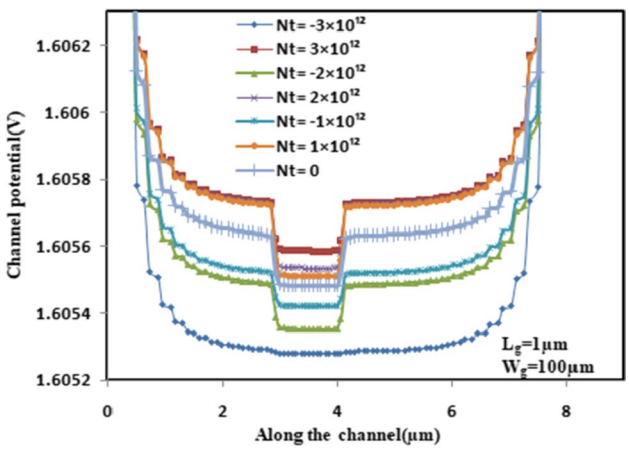


Fig. 4 Variation of channel potential along the channel for different radiation doses induced.

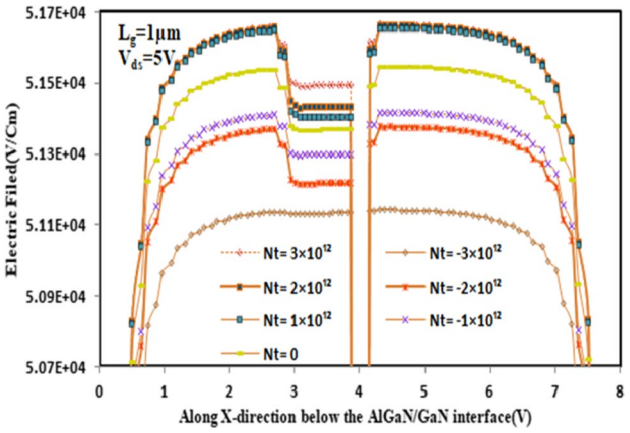


Fig. 5 Electric field plot along the *x*-direction of MOS-HEMT based dosimeter for different radiation dose induced.

V/cm and 5.11*E*+04 V/cm for $N_t = +3 \times 10^{12}\text{cm}^{-2}$ and $-3 \times 10^{12}\text{cm}^{-2}$, respectively, can be seen.

The contour plot for electron concentrations of different radiation doses induced is shown in Fig. 6. It has been observed that high radiation doses induced show a high impact on the charge carrier concentrations.

The radiation dose induced can be calculated with a change in threshold voltage shift from the following relationship[24]:

$$N_t = \frac{Cox\Delta V_{th}}{q} \tag{1}$$

where Cox is accumulation capacitance, N_t is the density of radiation dose, ΔV_{th} is threshold voltage shift, and q is the electron charge.

The sensitivity of the proposed sensor depends upon the radiation-generated interface trap charge. In Fig. 7, the change in the potential has been plotted for negative and positive radiation doses induced. It has been observed that whether we increase the positive trap charge or the negative trap charge, the change in potential of the device increases, as shown in Fig. 7. A maximum shift in the potential of 20 µV has been seen with a radiation dose of $N_t = -3 \times 10^{12}$ cm^{-2}.

Figure 8 shows the $I_{\text{ds}}-V_{\text{gs}}$ characteristics for various radiation doses induced of the AlGaN/GaN MOS-HEMT-based dosimeter at $V_{\text{ds}} = 5$ V. The radiation dose of $N_t = +3 \times 10^{12}\text{cm}^{-2}$ shows a maximum drain current of 83.2 mA and a slight variationin drain current of 13 mA at the gate bias of – 4 V. Therefore, the drain current increases with the value of the positive interface charge density (N_t).

Figure 9 shows a decrease in the drain current with an increase in the gate length for considered radiation doses (positive or negative). At $V_{\text{ds}} = 5$ V, the maximum drain current of 75 mA has been noticed for positively induced trap charges and a minimum drain current of 41 mA for

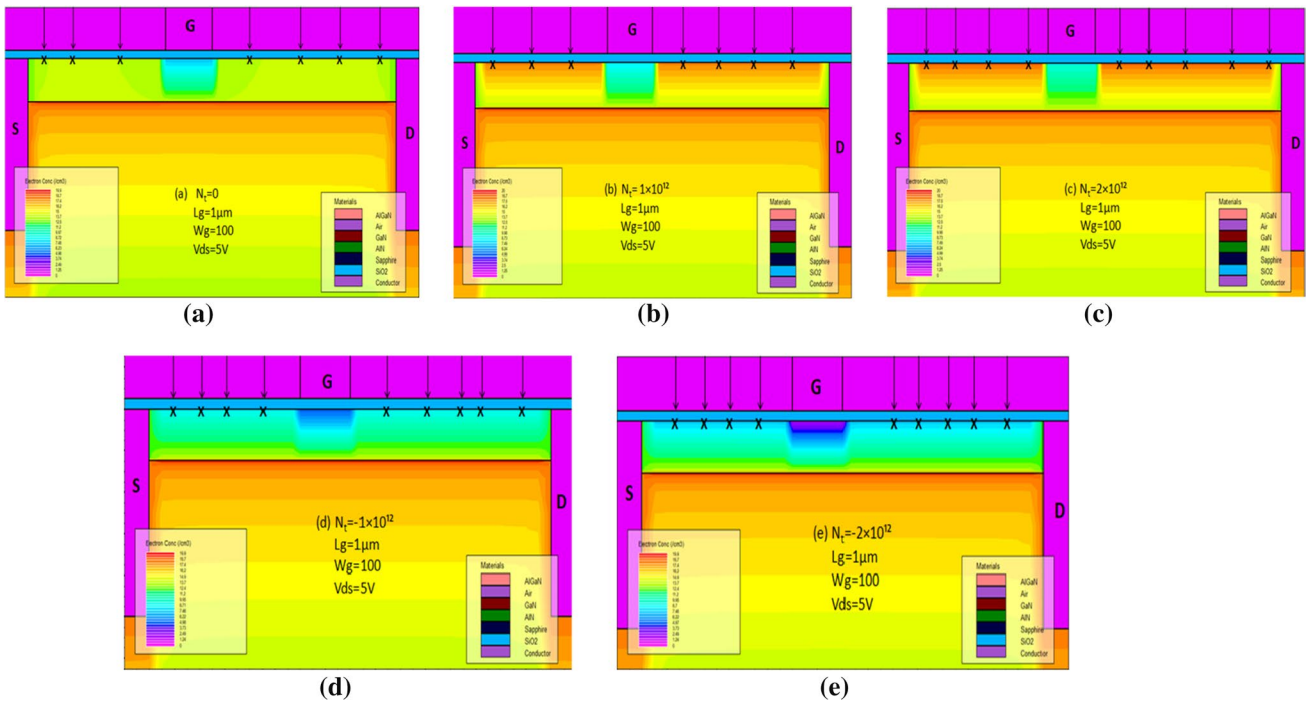


Fig. 6 Electron concentration contour plot of dosimeter-based MOS-HEMT (a) $Nt = 0 \times 10^{12}$ cm^{-2}, (b) $1 \times 10^{12}\text{cm}^{-2}$, (c) $2 \times 10^{12}\text{cm}^{-2}$, (d) $-1 \times 10^{12}\text{cm}^{-2}$, (e) $-2 \times 10^{12}\text{cm}^{-2}$.

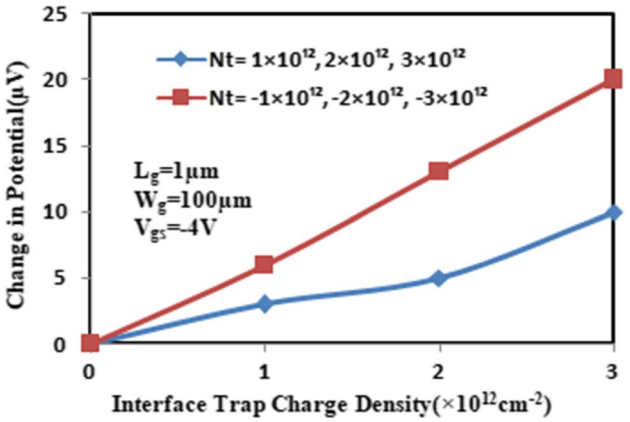


Fig. 7 Change in the potential of MOS-HEMT-based dosimeter for different radiation doses induced.

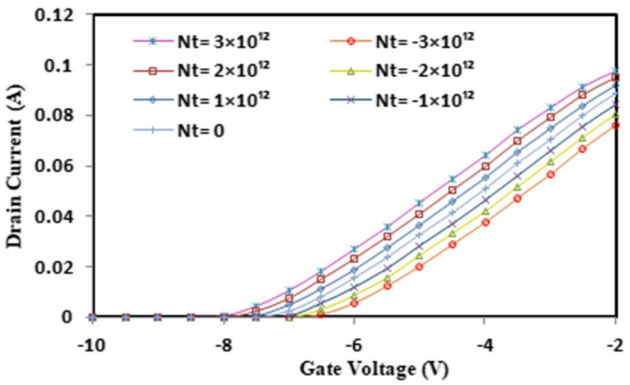


Fig. 8 $I_{ds}-V_{gs}$ characteristics of MOS-HEMT-based dosimeter.

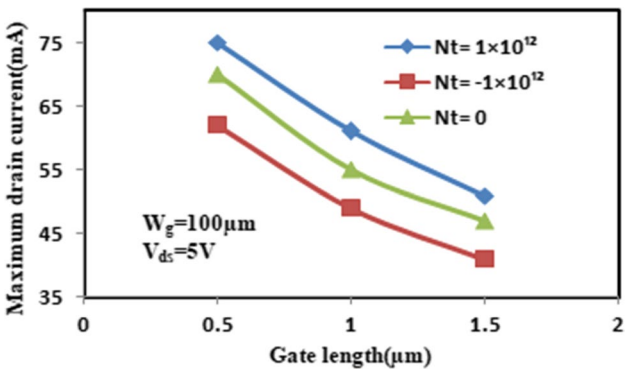


Fig. 9 Variation in drain current at different gate lengths (1, 0.5, 1.5 μm) for different radiation doses induced.

negatively induced trap charges. Because of the step-up in the gate length, the charge carriers have to travel long distances under the impact of gate bias, further reducing the free charge carriers in 2DEG, thus decreasing the drain current. Figure 10 shows the transconductance curve of the

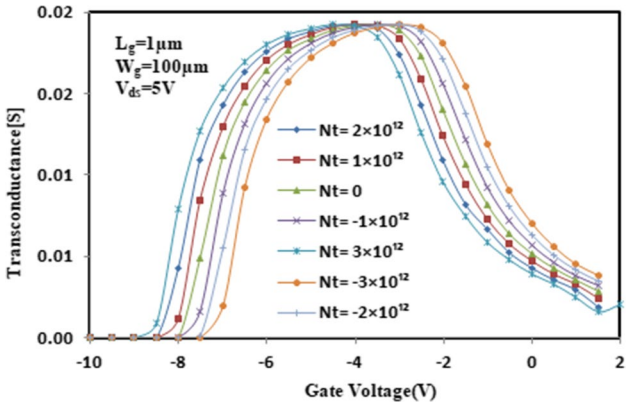


Fig. 10 Transconductance curve of MOS-HEMT-based induced dosimeter for different radiation doses

MOS-HEMT-based dosimeter for different radiation doses. Transconductance (g_m) for different radiation doses is shown in Table II. The radiation dose of $N_t = +3 \times 10^{12}\text{cm}^{-2}$offers a high transconductance of 0.0181S. At $N_t = +3 \times 10^{12}$ cm^{-2}, there is a high concentration of electrons in the 2DEG layer,increasing the device's drain current and threshold voltage.

Table II shows the comparative analysis of the device's positive and negative radiation dose-dependence on the performance of different characteristics (change in threshold voltage, sensitivity, transconductance, drain current). The maximum change in threshold voltage is 593 mV, and a maximum drain current of 0.073 A has been achieved in this device.

Figure 11 shows a threshold voltage shift for different radiation doses at different gate lengths (0.5, 1, 1.5 μm). The maximum threshold voltage of 7.053 V for the positive charge trap density was observed, and a minimum threshold voltage of 5.834 V for the negative charge trap density was monitored. From Fig. 11, it can also be seen that as the length of the gate increases, the threshold voltage of the device decreases.

$$\Delta V_{th} = \Delta V_{th}(WR) - \Delta V_{th}(WOR)\big|_{\text{at}I_{\text{ON}}=\text{constant}} \quad (2)$$

where $V_{th}(WR)$is the device threshold voltage with radiation when the radiation is considered on its surface, and $V_{th}(WOR)$ is the device threshold voltage without radiation when there is no radiation on the surface of the device.

The increase in either positive or negative radiation dose leads to a shift in the threshold voltage, and hence there will be a change in the sensitivity of the device. In Fig. 12, the threshold voltage has been plotted for different radiation doses with values $N_t = -3 \times 10^{12}$ cm^{-2} to $+3 \times 10^{12}$ cm^{-2}. An increase in the charge concentration results in a negative shift in the threshold voltage, because 2DEG in the

Table II Parameters for MOS-HEMT-based dosimeter device

Parameters	Radiation dose					
	-1×10^{12}cm^{-2}	-2×10^{12} cm^{-2}	-3×10^{12} cm^{-2}	1×10^{12} cm^{-2}	2×10^{12} cm^{-2}	3×10^{12}cm^{-2}
Sensitivity	0.092	0.185	0.277	0.129	0.222	0.314
Transconductance (S)	1.24*E*−02	1.09*E*−02	9.62*E*−03	1.57*E*−02	1.71*E*−02	1.81*E*−02
Change in threshold voltage (ΔV_{th}, mV)	203	482	651	182	488	593
Drain current (A)	0.050	0.045	0.039	0.061	0.067	0.073

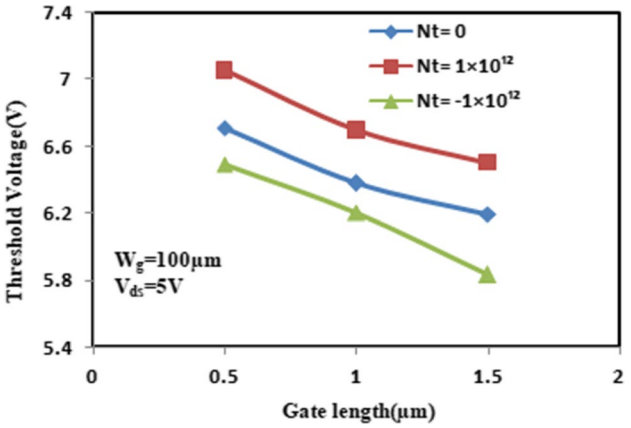


Fig. 11 Threshold voltage plot for gate lengths (0.5, 1, 1.5 µm) for different radiation doses induced.

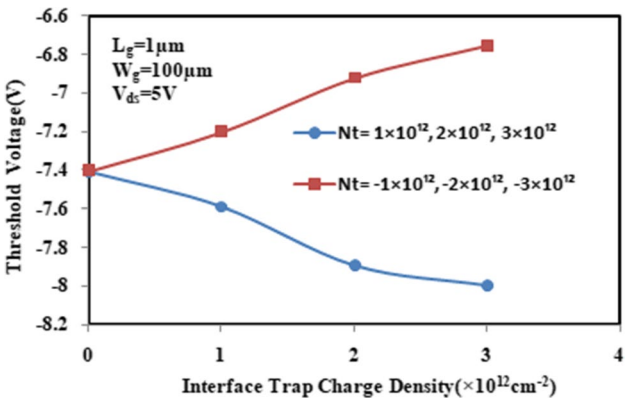


Fig. 12 Variations of the threshold voltage of MOS-HEMT dosimeter for different radiation doses induced.

channel is directly proportional to the threshold voltage. The proposed MOS-HEMT radiation-based AlGaN/GaN dosimeter indicates the minimum threshold voltage for a negative radiation dose because the charge carrier concentration at $N_t = +3\times10^{12}$cm^{-2} in 2DEG is higher. This result suggests that the threshold voltage increases with the radiation dose. A maximum threshold voltage of – 8 V was observed for the $N_t = -3\times10^{12}$cm^{-2} radiation dose.

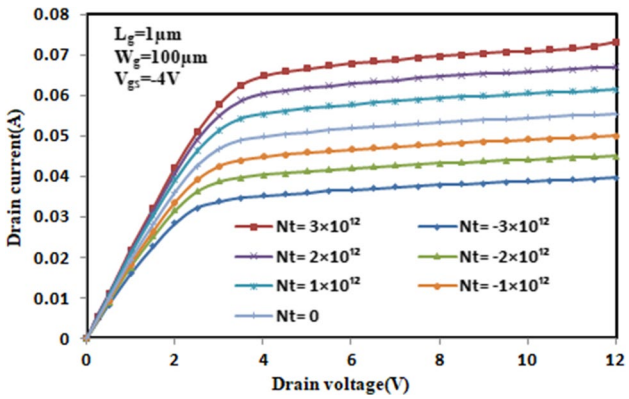


Fig. 13 $I_{ds}-V_{ds}$characteristics of MOS-HEMT dosimeter.

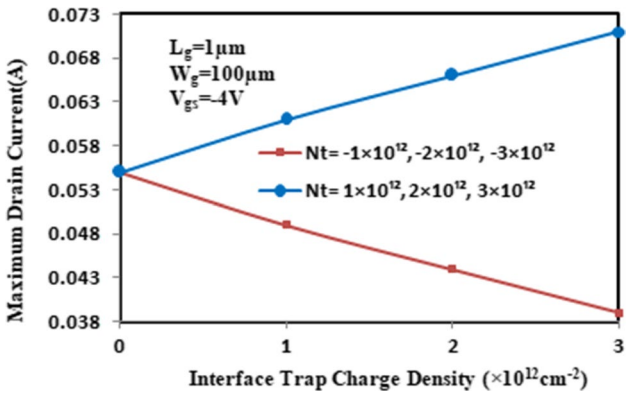


Fig. 14 Variation of drain current of MOS-HEMT based on different radiation doses induced.

Figure 13 shows the $I_{ds}-V_{ds}$characteristics for various radiation doses. The maximum drain current has been noticed for positive radiation doses induced. With the increase in the concentration of charge trap density, the drain current increases. A minimum current of 3.96*E*−02 A has been noticed at a drain voltage of 12 V. The charge carrier concentration (n_c) at vAlGaN/GaN interface[25] is given by:

$$n_c = \frac{\varepsilon_{\mathrm{AlGaN}}\left(\mathrm{V}_{gs} - \mathrm{V}_{\mathrm{off}}\right) - \mathrm{E}_f}{q \times \mathrm{t}_{\mathrm{AlGaN}}} \tag{3}$$

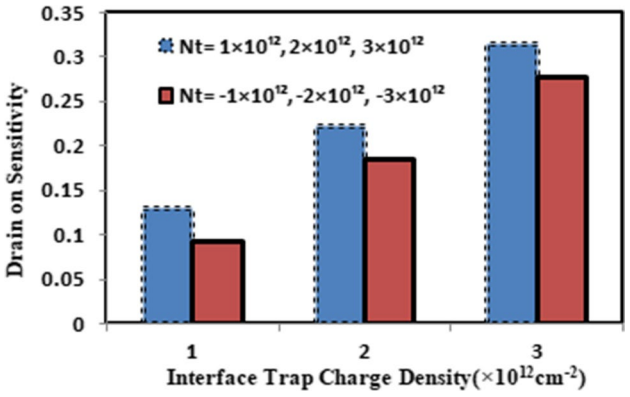


Fig. 15 Drain on the sensitivity of a MOS-HEMT-based dosimeter.

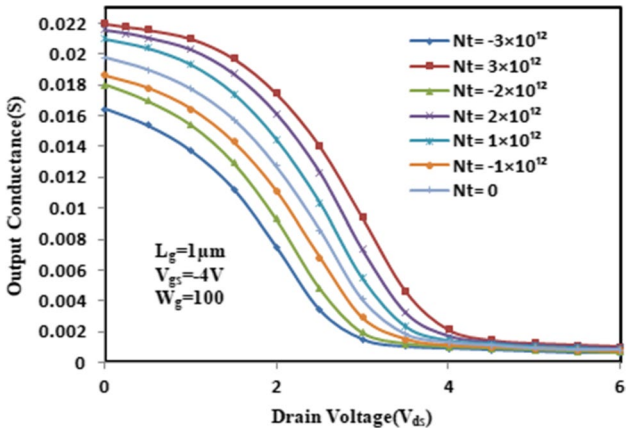


Fig. 16 Output conductance (gd) versus drain voltage for different radiation doses induced.

where $V_{off} = -4$ V is the cutoff voltage and V_{gs} is the applied gate voltage, ε_{AlGaN} is the permittivity of the barrier layer, E_f is the fermi level, and t_{AlGaN} is the thickness of the barrier layer. In Fig. 14, the maximum drain current is plotted for AlGaN/GaN MOS-HEMT for different radiation doses. The results show that, with an increase in the amount of positive radiation, the change in drain current increases; with an increase in the negative radiation dose, the change in the drain current decreases.

To maintain a charge neutrality condition in the device, equal amounts of charge concentration are increased in the 2DEG layer. The drain current increases for the charge density of the positive interface and decreases for the negative interface charge density. A maximum drain current of 0.071 A has been observed for $N_t = +3 \times 10^{12}\text{cm}^{-2}$, and a minimum value of 0.039 A has been observed for $N_t = -3 \times 10^{12}\text{cm}^{-2}$ at a drain voltage of 5 V.

The drain on snsitivity for a MOS-HEMT dosimeter is defined as[26]:

$$S = \left| \frac{I_{on(\text{without radiation})} - I_{on(\text{with radiation})}}{I_{on(\text{with radiation})}} \right|_{v_g=\text{Const.}} \quad (4)$$

where I_{on} (with radiation) is the drain current of the ON state of the device when radiation is considered on its surface, and I_{on} (without radiation) is the drain current of the ON state of the device when there is no radiation on its surface. Figure 15 shows the change in sensitivity parameters for various radiations doses induced. The increase in positive or negative radiation dose increases the drain on sensitivity. A higher positive charge results in the rise of the 2DEG charge, leading to an increased change in the drain current. Hence, the sensitivity is increased. Among the entire radiation doses, maximum and minimum sensitivities of 0.314, 0.092 for $N_t = +3 \times 10^{12}\text{cm}^{-2}$ and $-3 \times 10^{12}\text{cm}^{-2}$, respectively, have been observed because of the higher value of the trap charges taken into account in this device. Therefore, it is expected that the proposed dosimeter will show increased sensitivity for the high doses of radiation.

Figure16 shows the plot for output conductance (g_d) versus drain voltage with the considered charge concentrations. It can be seen from the figure that the highest change in output conductance can be seen at $V_{ds} = 2$ V. With an increase in the radiation dose induced, output conductance increases and, at higher drain bias, a change in the output conductance of 69% has been observed for $N_t = -3 \times 10^{12}\text{cm}^{-2}$.

Conclusions

A radiation-sensitive AlGaN/GaN MOS-HEMT-based dosimeter is proposed in this paper. Different radiation doses (N_t), $-3 \times 10^{12}\text{cm}^{-2}$ to $+3 \times 10^{12}\text{cm}^{-2}$, have been exposed to this device in terms of interface trap charges. Thus, it was found that variations in the sensitivity were detected by changing the radiation dose (N_t). The less the radiation dose (positive or negative), the less the threshold voltage shift. The electric field variations, threshold voltage, drain current, and surface potential have been calculated for different uniform radiation doses (N_t). The impact of the gate length on the drain current and threshold voltage sensitivity of the sensor has also been studied. The maximum change in output conductance, drain current, change in threshold voltage, drain on sensitivity, and transconductance noticed for the proposed sensor are 69%, 0.073 A, 593 mV, 0.314, and 0.018 S, respectively. It has also been found that our proposed MOS-HEMT-based dosimeter is efficient for radiation detection and hence can be employed as a dosimeter. To validate the results in terms of structure, the device was calibrated using Hu et al.'s[18] experimental results, and it shows a good match with the experimental data.

Acknowledgments The authors would like to thank the Director, Maharaja Agrasen Institute of Technology, Delhi, for carrying out this research work. This work has been carried out at the Device Research Lab, Maharaja Agrasen Institute of Technology, New Delhi, India.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. R. Dutta, T. Tamang, P. Paul, and N. Paitya, in *2020 4th Int. Conf. Trends Electron. Informatics* (IEEE, 2020), pp. 137–142.
2. A.-C. Liu, P.-T. Tu, C. Langpoklakpam, Y.-W. Huang, Y.-T. Chang, A.-J. Tzou, L.-H. Hsu, C.-H. Lin, H.-C. Kuo, and E.Y. Chang, *Micromachines* 12, 737 (2021).
3. A.S.A. Fletcher and D. Nirmal, *Superlattice Microst* 109, 519 (2017).
4. M.A.H. Khan, M.V. Rao, and Q. Li, *Sensors* 19, 905 (2019).
5. R.D. Schrimpf, in *Radiat. Eff. Embed. Syst.* (Springer, 2007), pp. 11–29.
6. D. Jena, E. Mohapatra, F. A. Ali, and T. P. Dash, in *2021 Devices Integr. Circuit* (IEEE, 2021), pp. 411–415.
7. K.T. Upadhyay and M.K. Chattopadhyay, *Mater. Sci. Eng. B* 263, 114849 (2021).
8. M.A.H. Khan and M.V. Rao, *Sensors* 20, 3889 (2020).
9. A. Meddour, F. Meddour, and S. Drid, Silicon 1 (2021).
10. S. L. Noor, S. Safa, and M. Z. R. Khan, in *2016 9th Int. Conf. Electr. Comput. Eng.* (IEEE, 2016), pp. 170–173.
11. L.S. August, *IEEE Trans. Nucl. Sci.* 31, 801 (1984).
12. P. Wan, J. Yang, T. Ying, G. Lv, L. Lva, S. Dong, L. Dongb, X. Yu, Z. Zhen, and W. Li, IEEE Trans. Nucl. Sci. (2021).
13. J.D. Howgate, M. Hofstetter, S.J. Schoell, M. Schmid, S. Schäfer, I. Zizak, V. Hable, C. Greubel, G. Dollinger, and S. Thalhammer, *Phys. Status Solidi* 209, 1562 (2012).
14. M. Soubra, J. Cygler, and G. Mackay, *Med. Phys.* 21, 567 (1994).
15. A. Malik, C. Sharma, R. Laishram, R.K. Bag, D.S. Rawal, S. Vinayak, and R.K. Sharma, *Solid. State. Electron.* 142, 8 (2018).
16. Z. Guo and T.P. Chow, *Phys. Status Solidi* 212, 1137 (2015).
17. A. Varghese, A. Eblabla, and K. Elgaid, *IEEE Sens. J.* 21, 15361 (2021).
18. T. Cheung, M.J. Butson, and K.N. Peter, *Phys. Med. Biol.* 49, N191 (2004).
19. P. Pal, Y. Pratap, M. Gupta, S. Kabra, and H. D. Sehgal, in *2020 5th Int. Conf. Devices, Circuits Syst.* (IEEE, 2020), pp. 203–206.
20. A. Dubey, A. Singh, R. Narang, M. Saxena, and M. Gupta, *IEEE Trans. Nanotechnol.* 17, 49 (2017).
21. C.-Y. Liu, Y.-C. Zhang, S.-R. Xu, L. Jiang, J.-C. Zhang, and Y. Hao, *Materials (Basel).* 12, 4050 (2019).
22. W.D. Hu, X.S. Chen, Z.J. Quan, X.M. Zhang, Y. Huang, C.S. Xia, W. Lu, and P.D. Ye, *J. Appl. Phys.* 102, 34502 (2007).
23. A.C.V. Bôas, M.A.A. de Melo, R.B.B. Santos, R. Giacomini, N.H. Medina, L.E. Seixas, S. Finco, F.R. Palomo, A. Romero-Maestre, and M.A. Guazzelli, *Microelectron. Reliab.* 116, 114000 (2021).
24. J. Zhu, X. Ma, B. Hou, M. Ma, Q. Zhu, L. Chen, L. Yang, P. Zhang, X. Zhou, and Y. Hao, *IEEE Trans. Electron Devices* 65, 5343 (2018).
25. F.M. Yigletu, S. Khandelwal, T.A. Fjeldly, and B. Iniguez, *IEEE Trans. Electron Devices* 60, 3746 (2013).
26. P. Pal, Y. Pratap, M. Gupta, and S. Kabra, Superlattices Microstruct. 106968 (2021).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

arXiv:2208.01241v1 [math.CV] 2 Aug 2022

RADIUS CONSTANTS OF SIGMOID STARLIKE FUNCTIONS

PRIYANKA GOEL AND S. SIVAPRASAD KUMAR*

Abstract. In the present investigation, we study the class of Sigmoid starlike functions, given by $\mathcal{S}^*_{SG} = \{f \in \mathcal{A} : zf'(z)/f(z) \prec 2/(1+e^{-z})\}$ in context of estimating the sharp radius constants associated with several known subclasses of starlike functions. Further, graphical validation for the sharpness of results is also provided.

1. Introduction

Let $\mathcal{A}_n$ be the class of analytic functions defined on $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$, satisfying $f(0) = 0,\ f'(0) = 1$ and of the form

$$f(z) = z + a_{n+1}z^{n+1} + a_{n+2}z^{n+2} + \cdots, \quad n \in \mathbb{N}.$$

We denote by $\mathcal{A} := \mathcal{A}_1$ and let $\mathcal{S} \subset \mathcal{A}$ be the class of univalent functions. For the functions f and F, we say that f is subordinate to F, written as $f \prec F$, if it is possible to write $f(z) = F(\omega(z))$ for some Schwarz function ω. Let $\mathcal{S}^*$ and $\mathcal{C}$ denote respectively the class of starlike and convex functions. Note that if we consider $f \in \mathcal{A}_n$, the class of starlike and convex functions are denoted by $\mathcal{S}^*_n$ and $\mathcal{C}_n$ respectively. For $n \in \mathbb{N}$, we now define the Carathéodory class $\mathcal{P}_n$, containing functions of the form $p(z) = 1 + c_n z^n + c_{n+1}z^{n+1} + \cdots$ with $\mathrm{Re}(p(z)) > 0$ on $\mathbb{D}$. Using subordination, Ma and Minda [11] defined a general subclass of starlike functions, given by $\mathcal{S}^*(\phi) := \{f \in \mathcal{A} : zf'(z)/f(z) \prec \phi(z)\}$. For different choices of ϕ, authors have defined several subclasses of $\mathcal{S}^*$ and examined these classes for different geometric properties. These classes will be described in the text wherever needed. In 2020, we introduced the class of Sigmoid starlike functions by taking $\phi(z) = 2/(1+e^{-z})$, the Modified Sigmoid function and denote it by $\mathcal{S}^*_{SG}$ [3]. The image of $\mathbb{D}$ under the Modified Sigmoid function is denoted by $\Delta_{SG} := \{w \in \mathbb{C} : |\log(w/(2-w))| < 1\}$. In [3], we present some basic geometry of this function, prove several inclusion relationships, obtain some coefficient bounds, and study mainly first-order differential subordination results. Later in [8], we proved various second and third order differential subordination results for Sigmoid starlike functions. Soon the class gained popularity and attracted many authors to study further in context of various aspects such as coefficient problems and convolution results (see [6, 13]). In continuation of these works, we now investigate $\mathcal{S}^*_{SG}$ for radius problems. We present in this paper, radius estimates for $\mathcal{S}^*_{SG}$ in conjunction with a bunch of other subclasses of starlike functions. Further, we consider certain families of analytic functions, which are characterized by the ratio of its functions with a specific function g and obtain $\mathcal{S}^*_{SG}$- radius for these families. We extensively use the following lemma in order to prove our main results:

Lemma 1.1. [3] *Let $2/(1+e) < a < 2e/(1+e)$. If*

$$r_a = \frac{e-1}{e+1} - |a-1|,$$

2020 *Mathematics Subject Classification.* 30C45, 30C80.

Key words and phrases. Radius problem, Sigmoid function, Starlike functions.

*Corresponding Author

The first author is supported by The Council of Scientific and Industrial Research(CSIR). Ref.No.:08/133(0018)/2017-EMR-I..

then

$$\{w \in \mathbb{C} : |w-a| < r_a\} \subset \Delta_{SG}. \tag{1.1}$$

2. Main Results

To begin with, we consider the class $\mathcal{S}_n^*[A,B]$ $(-1 \leq B < A \leq 1)$, introduced by Janowski [4], as below.

$$\mathcal{S}_n^*[A,B] := \left\{ f \in \mathcal{A}_n : \frac{zf'(z)}{f(z)} \prec \frac{1+Az}{1+Bz} \right\}.$$

Corresponding to the above class, we consider $\mathcal{P}_n[A,B]$ and a few results related to this class, given as follows:

$$\mathcal{P}_n^*[A,B] := \left\{ f \in \mathcal{P}_n : p(z) \prec \frac{1+Az}{1+Bz} \right\}.$$

Lemma 2.1. [15, Lemma 2.1]

(*i*) *If $p \in \mathcal{P}_n[A,B]$, then for $|z| = r$,*

$$\left| p(z) - \frac{1-ABr^{2n}}{1-B^2r^{2n}} \right| \leq \frac{(A-B)r^n}{1-B^2r^{2n}}.$$

(*ii*) *In particular, if $p \in \mathcal{P}_n(\alpha) := \mathcal{P}[1-2\alpha, -1]$, then*

$$\left| p(z) - \frac{1+(1-2\alpha)r^{2n}}{1-r^{2n}} \right| \leq \frac{2(1-\alpha)r^n}{1-r^{2n}}.$$

Lemma 2.2. [17, Lemma 2] *If $p \in \mathcal{P}_n(\alpha)$, then for $|z| = r$,*

$$\left| \frac{zp'(z)}{p(z)} \right| \leq \frac{2nr^n(1-\alpha)}{(1-r^n)(1+(1-2\alpha)r^n)}.$$

Theorem 2.3. *The sharp $\mathcal{S}_{SG,n}^*$-radius of the class $\mathcal{S}_n^*[A,B]$ is given by*

(*i*) $R_{\mathcal{S}_{SG,n}^*}(\mathcal{S}_n^*[A,B]) = \min\left\{1, \left(\frac{e-1}{A(1+e)-2B}\right)^{\frac{1}{n}}\right\}$, *when* $0 \leq B < A \leq 1$.

(*ii*) $R_{\mathcal{S}_{SG,n}^*}(\mathcal{S}_n^*[A,B]) = \min\left\{1, \left(\frac{e-1}{A(1+e)-2Be}\right)^{\frac{1}{n}}\right\}$, *when* $-1 \leq B < A \leq 1$ *with* $B \leq 0$.

In particular, for the class $\mathcal{S}^*$, we have $R_{\mathcal{S}_{SG}^*}(\mathcal{S}^*) = (e-1)/(3e+1)$.

Proof. Let $f \in \mathcal{S}_n^*[A,B]$. Using Lemma 2.1, we have

$$\left| \frac{zf'(z)}{f(z)} - \frac{1-ABr^{2n}}{1-B^2r^{2n}} \right| \leq \frac{(A-B)r^n}{1-B^2r^{2n}}. \tag{2.1}$$

(i) If $0 \leq B < A \leq 1$, then

$$a := \frac{1-ABr^{2n}}{1-B^2r^{2n}} \leq 1.$$

Further, we observe that $f \in \mathcal{S}_{SG,n}^*$ if the disk (2.1) is contained in the disk (1.1). Thus by using Lemma 1.1, it suffices to show that

$$\frac{(A-B)r^n}{1-B^2r^{2n}} \leq \frac{1-ABr^{2n}}{1-B^2r^{2n}} - \frac{2}{1+e},$$

which upon simplification, yields

$$r \leq \left(\frac{e-1}{A(1+e)-2B}\right)^{\frac{1}{n}}.$$

(ii) If $-1 \leq B < 0 < A \leq 1$, then

$$a := \frac{1 - ABr^{2n}}{1 - B^2 r^{2n}} \geq 1.$$

Again by Lemma 1.1, we see that $f \in \mathcal{S}^*_{SG,n}$ if

$$\frac{(A-B)r^n}{1 - B^2 r^{2n}} \leq \frac{2e}{1+e} - \frac{1 - ABr^{2n}}{1 - B^2 r^{2n}},$$

which simplifies to

$$r \leq \left(\frac{e-1}{A(1+e) - 2Be}\right)^{1/n}.$$

The result follows with sharpness due to the function $f_{A,B}(z)$, given by

$$f_{A,B}(z) = \begin{cases} z(1 + Bz^n)^{\frac{A-B}{nB}}; & B \neq 0, \\ z \exp\left(\frac{Az^n}{n}\right); & B = 0. \end{cases}$$

∎

Corollary 2.4. *The sharp $\mathcal{S}^*_{SG}$-radius for $\mathcal{S}^*(\alpha)$ is $(e-1)/(1+3e-2\alpha(1+e))$, $0 \leq \alpha < 1$. The bound is sharp for $k_\alpha(z) = z/(1-z)^{2(1-\alpha)}$.*

Corollary 2.5. *The sharp $\mathcal{S}^*_{SG}$-radius for $\mathcal{S}^*$ is $(e-1)/(1+3e)$. The bound is sharp for $k(z) = z/(1-z)^2$.*

Before we proceed to our next result, we need to recall the following classes:
For $0 \leq \alpha < 1$, the class $\mathcal{BS}^*(\alpha) := \{f \in \mathcal{A} : zf'(z)/f(z) \prec 1 + z/(1-\alpha z^2)\}$, defined by Kargar et al. [5]. In [7], Khatter et al. generalized $\mathcal{S}^*_L := \mathcal{S}^*(\sqrt{1+z})$ and $\mathcal{S}^*_e := \mathcal{S}^*(e^z)$ to $\mathcal{S}^*_L(\alpha) := \mathcal{S}^*(\alpha + (1-\alpha)\sqrt{1+z})$ and $\mathcal{S}^*_{\alpha,e} := \mathcal{S}^*(\alpha + (1-\alpha)e^z)$ respectively, for $\alpha \in [0,1)$.

Theorem 2.6. *The radius estimates of Sigmoid starlikeness, for the classes $\mathcal{BS}^*(\alpha)$, $\mathcal{S}^*_L(\alpha)$ and $\mathcal{S}^*_{\alpha,e}$ are given by*

(i) $R_{\mathcal{S}^*_{SG}}(\mathcal{BS}^*(\alpha)) = r_{\mathcal{BS}}(\alpha) := 2(e-1)/((1+e) + \sqrt{(1+e)^2 + 4\alpha(e-1)^2})$, *where* $\alpha \in [0,1)$.

(ii) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_L(\alpha)) = r_L(\alpha) := ((e-1)(3+e-2\alpha(1+e)))/((1-\alpha)^2(1+e)^2)$, *where* $\alpha \in [0, (3+e)/2(1+e))$. *In particular,* $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_L) = ((e-1)(3+e))/(1+e)^2$.

(iii) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{\alpha,e}) = r_e(\alpha) := \log(2e - \alpha(1+e))/(1+e)(1-\alpha)$, *where* $\alpha \in [0, (e(1+e)-2e)/((1+e)(e-1)))$. *In particular,* $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_e) = \log(2e/(1+e))$.

All estimates are sharp.

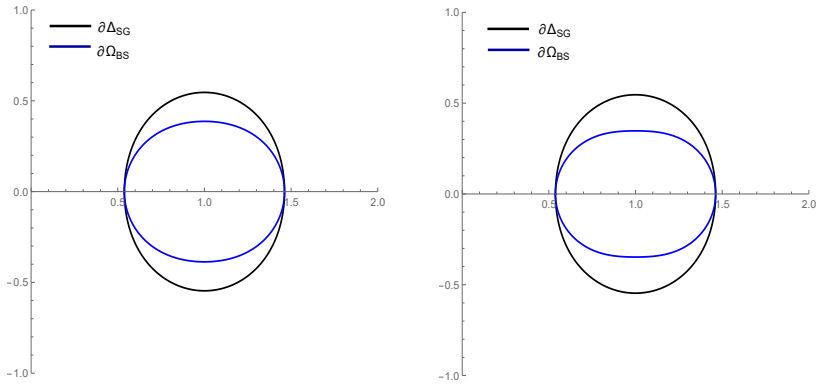
Proof. (i) Let $f \in \mathcal{BS}^*(\alpha)$. Then $zf'(z)/f(z) \prec 1 + z/(1-\alpha z^2)$ and thus

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \left|\frac{z}{1-\alpha z^2}\right| \leq \left|\frac{r}{1-\alpha r^2}\right| \quad \text{on } |z| = r.$$

Using Lemma 1.1, it can be said that the above disk lies in Δ_{SG} if $r/(1-\alpha r^2) \leq (e-1)/(e+1)$. This further implies $r \leq r_{\mathcal{BS}}(\alpha)$. Sharpness holds for the function

$$f_{\mathcal{BS}}(z) = \begin{cases} z\left(\frac{1+\sqrt{\alpha}z}{1-\sqrt{\alpha}z}\right)^{1/(2\sqrt{\alpha})}, & \alpha \in (0,1) \\ ze^z, & \alpha = 0. \end{cases}$$

It can be verified with the following graph that $zf'_{\mathcal{BS}}(z)/f_{\mathcal{BS}}(z)$ touches the boundary of Δ_{SG} at the points $\pm 2(e-1)/((1+e) + \sqrt{(1+e)^2 + 4\alpha(e-1)^2})$. Note that the domain Ω_{BS} denotes the image of $\mathbb{D}$ mapped by the function $1 + z/(1-\alpha z^2)$



Sharpness for $\alpha = 0.5$ Sharpness for $\alpha = 0.9$

FIGURE 1

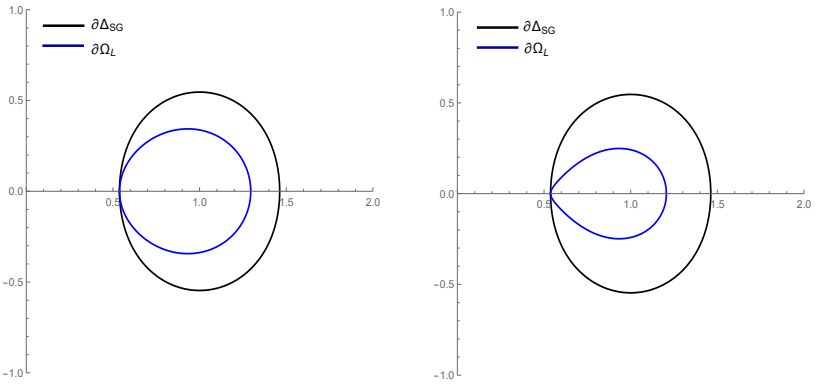
(ii) Let $f \in \mathcal{S}^*_L(\alpha)$, then $zf'(z)/f(z) \prec \alpha + (1-\alpha)\sqrt{1+z}$ and therefore on $|z|= r$

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq |(1-\alpha)(1-\sqrt{1+z})| \leq (1-\alpha)(1-\sqrt{1-r}).$$

By Lemma 1.1, it is clear that for the above disk to lie in Δ_{SG}, we need $(1-\alpha)(1-\sqrt{1-r}) \leq (e-1)/(e+1)$, which upon simplification yields $r \leq ((e-1)(3+e-2\alpha(1+e)))/((1-\alpha)^2(1+e)^2)$. Note that for the function

$$f_L(z) = z + (1-\alpha)z^2 + \frac{1}{16}(1-\alpha)(1-2\alpha)z^3 + \cdots,$$

the result is sharp. The sharpness of this result can be verified by the following graph, where Ω_L denotes the image of $\mathbb{D}$ mapped by $\alpha + (1-\alpha)\sqrt{1+z}$.



Sharpness for $\alpha = 0.1$ Sharpness for $\alpha = 0.5$

FIGURE 2

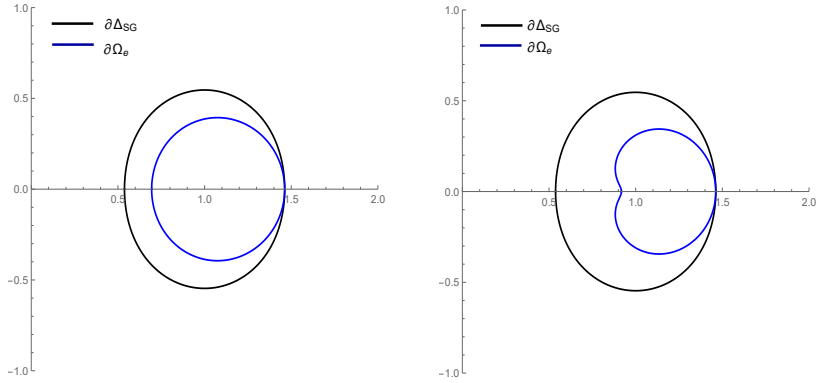
(iii) Let $f \in \mathcal{S}^*_{\alpha,e}$, then $zf'(z)/f(z) \prec \alpha + (1-\alpha)e^z$. So on $|z|= r$

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq (1-\alpha)|e^z - 1| \leq (1-\alpha)(e^r - 1).$$

By Lemma 1.1, $f \in \mathcal{S}^*_{SG}$ if $(1-\alpha)(e^r-1) \leq (e-1)/(e+1)$, which is equivalent to $r \leq r_e(\alpha)$. The result is sharp for the function

$$f_e(z) = z + (1-\alpha)z^2 + \frac{1}{4}(1-\alpha)(3-2\alpha)z^3 + \cdots$$

and is validated by the following graph. The image of $\mathbb{D}$ mapped by $\alpha + (1-\alpha)e^z$ is denoted by Ω_e.



Sharpness for $\alpha = 0.1$ Sharpness for $\alpha = 0.9$

FIGURE 3

■

Before we proceed further, let us recall the following classes: In [14], Mendiratta et al. considered the class of starlike functions associated with right lemniscate of Bernoulli, denoted by $\mathcal{S}^*_{RL} = \mathcal{S}^*(\phi)$, where ϕ is given by

$$\phi(z) = \sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1-z}{1+2(\sqrt{2}-1)z}}.$$

The class of cardioid starlike functions, denoted by $\mathcal{S}^*_C := \mathcal{S}^*(1+4z/3+2z^2/3)$, defined by Sharma et al. [18]. For $k = \sqrt{2}+1$, Kumar and Ravichandran [10] introduced $\mathcal{S}^*_R$ by taking ϕ as $1 + z(k+z)/k(k-z)$.

Theorem 2.7. *The sharp $\mathcal{S}^*_{SG}$-radius for the classes $\mathcal{S}^*_{RL}$, $\mathcal{S}^*_C$ and $\mathcal{S}^*_R$ is given by:*

(*i*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{RL}) =: r_{RL} = \left(4\sqrt{2}-7e-5\right)(e-1)/(32\sqrt{2}-7e^2+6e\left(4\sqrt{2}-5\right)-47) \approx 0.738309.$

(*ii*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_C) =: r_C = -1+\sqrt{(-1+5e)/(2+2e)} \approx 0.301221.$

(*iii*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_R) =: r_R = (\sqrt{\left(2\sqrt{2}+3\right)(2e^2-1)} - \left(\sqrt{2}+1\right)e)/(1+e) \approx 0.645131.$

Proof. (i) Let $f \in \mathcal{S}^*_{RL}$. Then

$$\frac{zf'(z)}{f(z)} \prec \sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1-z}{1+2(\sqrt{2}-1)z}}.$$

Thus on $|z|=r$, we have

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq 1 - \left(\sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1+r}{1-2(\sqrt{2}-1)r}}\right).$$

By Lemma 1.1, f is in $\mathcal{S}^*_{SG}$ if

$$1-\left(\sqrt{2}-(\sqrt{2}-1)\sqrt{\frac{1+r}{1-2(\sqrt{2}-1)r}}\right)\leq\frac{e-1}{e+1},$$

which is equivalent to $r\leq r_{RL}$. This result is sharp for the following function

$$f_{RL}(z)=z\left(\frac{\sqrt{1-z}+\sqrt{1+2(\sqrt{2}-1)z}}{2}\right)^{2\sqrt{2}-2}\exp\left(\sqrt{2(\sqrt{2}-1)}\tan^{-1}\Psi(z)\right),$$

where

$$\Psi(z)=\frac{\sqrt{2(\sqrt{2}-1)}\left(\sqrt{2(\sqrt{2}-1)z+1}-\sqrt{1-z}\right)}{2(\sqrt{2}-1)\sqrt{1-z}+\sqrt{2(\sqrt{2}-1)z+1}}.$$

The sharpness of this bound can be verified from Figure 4(i).

(ii) Let $f\in\mathcal{S}^*_C$, then we have $zf'(z)/f(z)\prec 1+4z/3+2z^2/3$. Therefore on $|z|=r$, we get

$$\left|\frac{zf'(z)}{f(z)}-1\right|\leq\frac{2(r^2+2r)}{3},$$

which if not exceeds $(e-1)/(e+1)$, implies that f lies in $\mathcal{S}^*_{SG}$, by Lemma 1.1. Solving this, we get $r\leq r_C$. In order to verify the sharpness of this result, we consider the following function.

$$f_C(z)=z\exp\left(\frac{4z}{3}+\frac{z^2}{3}\right).$$

Clearly $f_C\in\mathcal{S}^*_C$ and moreover $zf'_C(z)/f_C(z)$ touches the boundary of Δ_{SG} at the point $z_0=-1+\sqrt{(-1+5e/(2+2e)}$, as shown in Figure 4(ii).

(iii) Let $f\in\mathcal{S}^*_R$, then

$$\frac{zf'(z)}{f(z)}\prec 1+\frac{z(k+z)}{k(k-z)},$$

where $k=\sqrt{2}+1$. Thus on $|z|=r$,

$$\left|\frac{zf'(z)}{f(z)}-1\right|\leq\frac{r(k+r)}{k(k-r)}.$$

In view of Lemma 1.1, $f\in\mathcal{S}^*_{SG}$ if $r(k+r)/k(k-r)\leq(e-1)/(e+1)$. Solving this inequality, we obtain $r\leq r_R$. The equality of the radius estimate holds for the function

$$f_R(z)=\frac{k^2z}{(k-z)^2}e^{-z/k},\quad k=\sqrt{2}+1.$$

Figure 4(iii) verifies the sharpness of the result. Note that Ω_{RL}, Ω_c and Ω_R denote the image of $\mathbb{D}$ mapped by $zf'(z)/f(z)$ for f_{RL}, f_C and f_R repectively.

∎

Let us recall the following classes in order to obtain our next result. By taking $\phi(z)=z+\sqrt{1+z^2}$, Sharma et al. [19] introduced $\mathcal{S}^*_{\mathbb{C}}$. Similarly, Kumar and Gangania [9] intoduced $\mathcal{S}^*_{\wp}$ by taking ϕ as $1+ze^z$, the cardioid function.

Theorem 2.8. *The sharp $\mathcal{S}^*_{SG}-$radii for the classes $\mathcal{S}^*_{\mathbb{C}}$ and $\mathcal{S}^*_{\wp}$ is given by*

(*i*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{\mathbb{C}})=\frac{-1-2e+3e^2}{4e+4e^2}\approx 0.389089.$

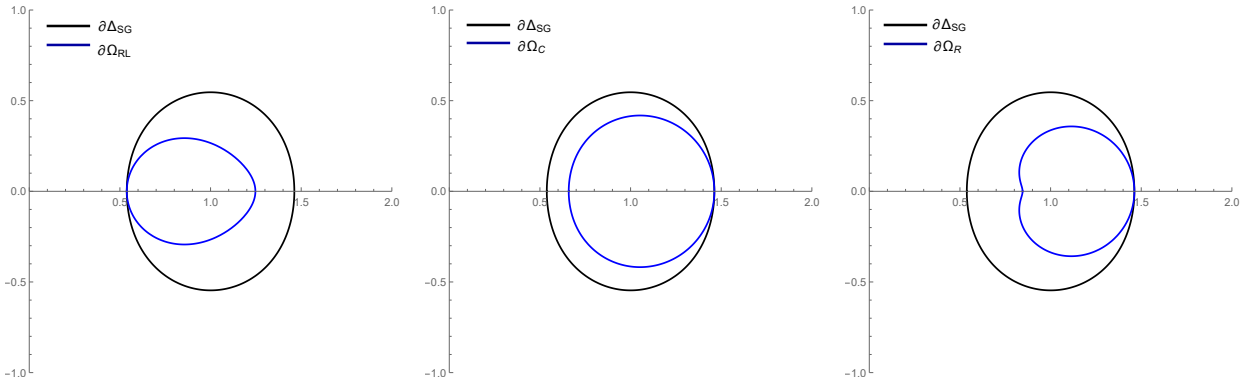


FIGURE 4

(*ii*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{\wp}) = r_{\wp} \approx 0.331672$, *where* $r_{\wp}$ *is the smallest positive root of the equation* $(e+1)re^r = e-1$.

Proof. (*i*) Let $f \in \mathcal{S}^*_{\mathbb{C}}$, then $zf'(z)/f(z) \prec z + \sqrt{1+z^2}$. Therefore on $|z| = r$,

$$\left|\frac{zf'(z)}{f(z)} - 1\right| = |z + \sqrt{1+z^2} - 1| \leq r + \sqrt{1+r^2} - 1.$$

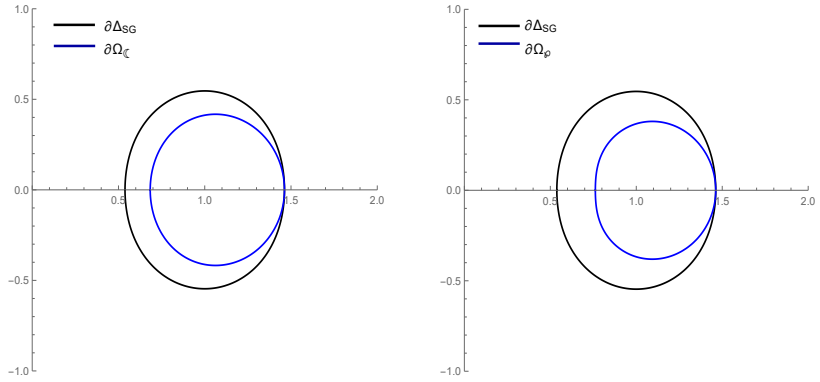
Now by using Lemma 1.1, the above disk lies inside the domain Δ_{SG} if $r + \sqrt{1+r^2} - 1 \leq (e-1)/(e+1)$. Solving this equation, we obtain the desired bound of r. The result is sharp for the function

$$f_{\mathbb{C}}(z) = z \exp\left(\int_0^z \frac{t + \sqrt{1+t^2} - 1}{t} dt\right) = z + z^2 + \frac{3z^3}{4} + \frac{5z^4}{12} + \frac{z^5}{6} + \cdots.$$

(*ii*) Let $f \in \mathcal{S}^*_{\wp}$, then it is clear that $zf'(z)/f(z) \prec 1 + ze^z$. So we have

$$\left|\frac{zf'(z)}{f(z)} - 1\right| = |ze^z| \leq re^r \quad \text{on } |z| = r.$$

By using Lemma 1.1, we can say that $f \in \mathcal{S}^*_{SG}$ if $re^r \leq (e-1)/(e+1)$, which is equivalent to $r \leq r_{\wp}$. The sharpness of the result can be verified by the function $f_{\wp}(z) = ze^{e^z - 1}$. The following graph depicts the sharpness of both the estimates. Note that the image of $\mathbb{D}$ mapped by $z + \sqrt{1+z^2}$ and $1 + ze^z$ are respectively denoted by $\Omega_{\mathbb{C}}$ and $\Omega_{\wp}$.



Sharpness of $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{\mathbb{C}})$ Sharpness of $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{\wp})$

FIGURE 5

∎

Now we consider the following classes for our next result. The class $\mathcal{S}^*_{Ne}$, defined by Wani and Swaminathan [21] by taking ϕ as $1+z-z^3/3$ and the class $\mathcal{S}^*_S = \mathcal{S}^*(1+\sin z)$, introduced by Cho et al. [2].

Theorem 2.9. *The $\mathcal{S}^*_{SG}$-radius for the classes $\mathcal{S}^*_{Ne}$ and $\mathcal{S}^*_S$ is given by:*

(*i*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_{Ne}) = r_{Ne} \approx 0.43473$, *which is the smallest positive root of the equation* $(e+1)(3r+r^3) = 3(e-1)$.

(*ii*) $R_{\mathcal{S}^*_{SG}}(\mathcal{S}^*_S) = \log\left(\frac{\sqrt{2(1+e^2)}+e-1}{1+e}\right) \approx 0.447074.$

Proof. (*i*) Let $f \in \mathcal{S}^*_{Ne}$, then $zf'(z)/f(z) \prec 1+z-z^3/3$. Thus on $|z|= r$,

$$\left|\frac{zf'(z)}{f(z)} - 1\right| = |z - \frac{z^3}{3}| \leq r + \frac{r^3}{3},$$

which if, not greater than $(e-1)/(e+1)$, is sufficient to conclude that $f \in \mathcal{S}^*_{SG}$ by Lemma 1.1. Note that for $r \leq r_{Ne}$, the above inequality holds and hence the result.

(*ii*) Let $f \in \mathcal{S}^*_S$, then $zf'(z)/f(z) \prec 1+\sin z$. So on $|z|= r$, we have

$$\left|\frac{zf'(z)}{f(z)} - 1\right| = |\sin z| \leq \sinh r.$$

Now by using Lemma 1.1, we observe that for f to belong to $\mathcal{S}^*_{SG}$, it is sufficient to prove that $\sinh r \leq (e-1)/(e+1)$. Solving this inequality, we obtain the desired result.

∎

Further let us consider the classes, introduced by Ali et al. [1], defined as follows:

$$\begin{aligned}
\mathcal{G}_1 &= \left\{f \in \mathcal{A}_n : \frac{f}{g} \in \mathcal{P}_n \text{ and } \frac{g(z)}{z} \in \mathcal{P}_n,\ g \in \mathcal{A}_n\right\}\\
\mathcal{G}_2 &= \left\{f \in \mathcal{A}_n : \frac{f}{g} \in \mathcal{P}_n \text{ and } \frac{g(z)}{z} \in \mathcal{P}_n(1/2),\ g \in \mathcal{A}_n\right\}\\
\mathcal{G}_3 &= \left\{f \in \mathcal{A}_n : \left|\frac{f(z)}{g(z)} - 1\right| < 1 \text{ and } \frac{g(z)}{z} \in \mathcal{P}_n,\ g \in \mathcal{A}_n\right\}\\
\mathcal{G}_4 &= \left\{f \in \mathcal{A}_n : \left|\frac{f(z)}{g(z)} - 1\right| < 1,\ g \in \mathcal{C}_n\right\}.
\end{aligned}$$

Theorem 2.10. *The $\mathcal{S}^*_{SG,n}$ radius of the class $\mathcal{G}_1$ is given by*

$$R_{\mathcal{S}^*_{SG,n}}(\mathcal{G}_1) = r_{\mathcal{G}_1} := \frac{(e-1)^{\frac{1}{n}}}{(2n(1+e) + \sqrt{4n^2(1+e)^2 + (e-1)^2})^{\frac{1}{n}}}.$$

The estimate is sharp.

Proof. Let $f \in \mathcal{G}_1$, then there exists $p,\ q \in \mathcal{P}_n$ such that

$$p(z) = \frac{g(z)}{z} \quad \text{and} \quad q(z) = \frac{f(z)}{g(z)}.$$

By Lemma 2.2, we have on $|z|= r$, the following inequalities

$$\left|\frac{zp'(z)}{p(z)}\right| \leq \frac{2nr^n}{1-r^{2n}} \quad \text{and} \quad \left|\frac{zq'(z)}{q(z)}\right| \leq \frac{2nr^n}{1-r^{2n}}. \tag{2.2}$$

Further, note that $f(z) = g(z)q(z) = zp(z)q(z)$, which upon logarithmic differentiation yields

$$\frac{zf'(z)}{f(z)} = 1 + \frac{zp'(z)}{p(z)} + \frac{zq'(z)}{q(z)}. \tag{2.3}$$

Using (2.2), we obtain

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{4nr^n}{1-r^{2n}}.$$

By Lemma 1.1, it is clear that $f \in \mathcal{S}^*_{SG}$ if the quantity $4nr^n/(1-r^{2n})$ does not exceed $(e-1)/(e+1)$. This leads us to $(e-1)r^{2n} + 4n(e+1)r^n - (e-1) \leq 0$. Solving this inequality, we obtain $r \leq r_{\mathcal{G}_1}$. The sharpness of the bound can be verified by the functions

$$f_1(z) = z\left(\frac{1+z^n}{1-z^n}\right)^2 \quad \text{and} \quad g_1(z) = z\left(\frac{1+z^n}{1-z^n}\right).$$

Clearly $f_1, g_1 \in \mathcal{A}_n$ and $f_1(z)/g_1(z) = g_1(z)/z$, which belongs to $\mathcal{P}_n$. So $f_1 \in \mathcal{G}_1$ and

$$\frac{zf_1'(z)}{f_1(z)} = 1 + \frac{4nz^n}{1-z^{2n}}.$$

The image domain of $\mathbb{D}$ mapped by $zf_1'(z)/f_1(z)$ touches the boundary of Δ_{SG} at the points $\pm r_{\mathcal{G}_1}$. ∎

Theorem 2.11. *The $\mathcal{S}^*_{SG,n}$ radius of the class $\mathcal{G}_2$ is given by*

$$R_{\mathcal{S}^*_{SG,n}}(\mathcal{G}_2) = r_{\mathcal{G}_2} := \frac{(2(e-1))^{\frac{1}{n}}}{(3n(1+e) + \sqrt{(3n(e+1))^2 + 4(e-1)(n(e+1)+(e-1))})^{\frac{1}{n}}}.$$

The estimate is sharp.

Proof. Let $f \in \mathcal{G}_2$. Now, let us define

$$p(z) = \frac{g(z)}{z} \quad \text{and} \quad q(z) = \frac{f(z)}{g(z)}$$

so that $f(z) = zp(z)q(z)$, where $p \in \mathcal{P}_n(1/2)$ and $q \in \mathcal{P}_n$. From (2.3) and Lemma 2.2, we have

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{2nr^n}{1-r^{2n}} + \frac{nr^n}{1-r^n} = \frac{nr^{2n} + 3nr^n}{1-r^{2n}} \quad \text{on } |z| = r.$$

In view of Lemma 1.1, f is a member of $\mathcal{S}^*_{SG}$ if $(nr^{2n} + 3nr^n)/(1-r^{2n}) \leq (e-1)/(e+1)$. Equivalently, $(n(e+1)+e-1)r^{2n} + 3n(e+1)r^n - (e-1) \leq 0$. Solving this inequality, we obtain $r \leq r_{\mathcal{G}_2}$. The sharpness of the bound can be verified by the functions

$$f_2(z) = \frac{z(1+z^n)}{(1-z^n)^2} \quad \text{and} \quad g_2(z) = \frac{z}{1-z^n}.$$

Clearly $f_2, g_2 \in \mathcal{A}_n$, $f_2/g_2 \in \mathcal{P}_n$ and $g_2/z \in \mathcal{P}(1/2)$. Hence f_2 is a member of $\mathcal{G}_2$ and

$$\frac{zf_2'(z)}{f_2(z)} = \frac{1 + 3nz^n + (n-1)z^{2n}}{1-z^{2n}}.$$

The radius estimate is sharp since $zf_2'(z)/f_2(z)$ maps $\mathbb{D}$ onto a domain which touches the boundary of Δ_{SG} at $z = r_{\mathcal{G}_2}$. ∎

Theorem 2.12. *The $\mathcal{S}^*_{SG,n}$ radius of the class $\mathcal{G}_3$ is given by*

$$R_{\mathcal{S}^*_{SG,n}}(\mathcal{G}_3) = r_{\mathcal{G}_3} := \frac{(2(e-1))^{\frac{1}{n}}}{(3n(1+e) + \sqrt{(3n(e+1))^2 + 4(e-1)(n(e+1)+(e-1))})^{\frac{1}{n}}}.$$

The estimate is sharp.

Proof. Let $f \in \mathcal{G}_3$. So, we may define

$$p(z) = \frac{g(z)}{z} \quad \text{and} \quad q(z) = \frac{g(z)}{f(z)}.$$

In view of the following implication

$$\left|\frac{f(z)}{g(z)} - 1\right| < 1 \iff \frac{g}{f} \in \mathcal{P}_n\left(\frac{1}{2}\right),$$

it is obvious that $p \in \mathcal{P}_n$ and $q \in \mathcal{P}_n(1/2)$. Note that $f(z) = zp(z)/q(z)$ and thus

$$\frac{zf'(z)}{f(z)} = 1 + \frac{p'(z)}{p(z)} - \frac{zq'(z)}{q(z)}.$$

By using Lemma 2.2, we obtain

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{nr^{2n} + 3nr^n}{1 - r^{2n}} \quad \text{on } |z| = r.$$

In order to prove $f \in \mathcal{S}^*_{SG}$, it suffices to show that $(nr^{2n} + 3nr^n)/(1 - r^{2n}) \leq (e-1)/(e+1)$, in view of Lemma 1.1. This inequality reduces to $(n(e+1) + e - 1)r^{2n} + 3n(e+1)r^n - (e-1) \leq 0$. Solving this, we obtain $r \leq r_{\mathcal{G}_3}$. The sharpness of the bound can be verified by the functions

$$f_3(z) = \frac{z(1+z^n)^2}{1-z^n} \quad \text{and} \quad g_3(z) = \frac{z(1+z^n)}{1-z^n}.$$

Note that

$$\left|\frac{f_3(z)}{g_3(z)} - 1\right| = |z|^n < 1 \quad \text{and} \quad \frac{g_3(z)}{z} = \frac{1+z^n}{1-z^n} \in \mathcal{P}_n,$$

which implies that $f_3 \in \mathcal{G}_3$. Moreover, it can be observed that the domain $(zf_3'/f_3)(\mathbb{D})$ touches the boundary of Δ_{SG} at $-r_{\mathcal{G}_3}$. ∎

Theorem 2.13. *The $\mathcal{S}^*_{SG,n}$ radius of the class $\mathcal{G}_4$ is given by*

$$R_{\mathcal{S}^*_{SG,n}}(\mathcal{G}_4) = r_{\mathcal{G}_4} := \frac{(2(e-1))^{\frac{1}{n}}}{((n+1)(1+e) + \sqrt{(n+1)^2(1+e)^2 + 4(e-1)((e+1)n-2)})^{\frac{1}{n}}}.$$

The estimate is sharp.

Proof. Let $f \in \mathcal{G}_4$. Now, suppose $q(z) = g(z)/f(z)$, where $g \in \mathcal{A}_n$ is a convex function. As deduced in the last theorem, we know that $q \in \mathcal{P}_n(1/2)$. By using Lemma 2.2, we obtain

$$\left|\frac{zq'(z)}{q(z)}\right| \leq \frac{nr^n}{1-r^n}, \quad |z| = r. \tag{2.4}$$

Also since g is convex, we have zg'/g is in $\mathcal{P}(1/2)$. Hence

$$\left|\frac{zg'(z)}{g(z)} - \frac{1}{1-r^{2n}}\right| \leq \frac{r^n}{1-r^{2n}} \tag{2.5}$$

by Lemma 2.2. Since $f = g/q$, we obtain

$$\frac{zf'(z)}{f(z)} = \frac{zg'(z)}{g(z)} - \frac{zq'(z)}{q(z)}.$$

In view of (2.4) and (2.5), we have

$$\left|\frac{zf'(z)}{f(z)} - \frac{1}{1-r^{2n}}\right| \leq \frac{(n+1)r^n + nr^{2n}}{1-r^{2n}}.$$

For each $r \in (0,1)$, the quantity $1-r^{2n}$ is less than 1 and thus Lemma 1.1 implies that $f \in \mathcal{S}^*_{SG}$ provided

$$\frac{(n+1)r^n + nr^{2n}}{1-r^{2n}} \leq \frac{1}{1-r^{2n}} - \frac{2}{1+e}.$$

This inequality reduces to

$$((1+e)n-2)r^{2n} + (n+1)(1+e)r^n + 1 - e \leq 0.$$

Solving this inequality, we obtain $r \leq r_{\mathcal{G}_4}$. The functions

$$f_4(z) = \frac{z(1+z^n)}{(1-z^n)^{1/n}} \quad \text{and} \quad g_4(z) = \frac{z}{(1-z^n)^{1/n}}$$

validate the sharpness of the bound. Note that $|f_4(z)/g_4(z) - 1| = |z|^n < 1$ and $g_4 \in \mathcal{C}_n$, which implies $f_4 \in \mathcal{G}_4$. Moreover, it can be observed that zf_4'/f_4 maps $\mathbb{D}$ onto a domain that touches the boundary of Δ_{SG} at $\pm r_{\mathcal{G}_4}$. ∎

For $0 \leq \alpha < 1$, Reade [16] introduced the class of close-to-starlike functions of type α, given by

$$\mathcal{CS}^*(\alpha) := \left\{ f \in \mathcal{A}_n : \frac{f}{g} \in \mathcal{P}_n,\ g \in \mathcal{S}^*(\alpha) \right\}.$$

Theorem 2.14. *The sharp $\mathcal{S}^*_{SG,n}$ radius for the class $\mathcal{CS}^*_n(\alpha)$ is given by*

$$\begin{aligned} R_{\mathcal{S}^*_{SG,n}}(\mathcal{CS}^*_n(\alpha)) = r_{cs} &:= (e-1)/((1+e)(1+n-\alpha) \\ &+ \sqrt{(e+1)^2(1+n-\alpha)^2 + (e-1)((1-2\alpha)(e+1)+2e)}) \end{aligned}$$

Proof. Let $f \in \mathcal{CS}^*_n(\alpha)$. Then there is some function $g \in \mathcal{S}^*_n(\alpha)$ such that $h = f/g \in \mathcal{P}_n$. Using Lemma 2.2, we get

$$\left| \frac{zh'(z)}{h(z)} \right| \leq \frac{2nr^n}{1-r^{2n}}. \tag{2.6}$$

Since $g \in \mathcal{S}^*_n(\alpha)$ and $zg'/g \in \mathcal{P}_n(\alpha)$, again by using Lemma 2.2, we obtain

$$\left| \frac{zg'(z)}{g(z)} - \frac{1+(1-2\alpha)r^{2n}}{1-r^{2n}} \right| \leq \frac{2(1-\alpha)r^n}{1-r^{2n}}. \tag{2.7}$$

Since $f = hg$, we have

$$\frac{zf'(z)}{f(z)} = \frac{zg'(z)}{g(z)} + \frac{zh'(z)}{h(z)}$$

From (2.6) and (2.7), it follows that

$$\left| \frac{zf'(z)}{f(z)} - \frac{1+(1-2\alpha)r^{2n}}{1-r^{2n}} \right| \leq \frac{2(1+n-\alpha)r^n}{1-r^{2n}}.$$

Note that the above inequality represents a disk with center a and radius r_0 given by

$$a := \frac{1+(1-2\alpha)r^{2n}}{1-r^{2n}} \quad \text{and} \quad r_0 := \frac{2(1+n-\alpha)r^n}{1-r^{2n}}.$$

Since $a > 1$, we have from Lemma 1.1 that f is a member of the class $\mathcal{S}^*_{SG}$ if $r_0 \leq (2e)/(e+1) - a$ or equivalently,

$$((1-2\alpha)(e+1)+2e)r^{2n} + 2(e+1)(1+n-\alpha)r^n + 1 - e \leq 0.$$

Solving this inequality, we obtain $r \leq r_{cs}$. The sharpness of this result can be verified by the following functions.

$$f(z) = \frac{z(1+z^n)}{(1-z)^{(n+2-2\alpha)/n}} \quad \text{and} \quad g(z) = \frac{z}{(1-z^n)^{(2-2\alpha)/n}}.$$

Note that $f/g = (1+z^n)/(1-z^n) \in \mathcal{P}_n$ and $g \in \mathcal{S}^*_n(\alpha)$, which ensures that $f \in \mathcal{CS}^*_n(\alpha)$. ∎

For our next result, let us recall the class $\mathcal{W}_n$, given as

$$\mathcal{W}_n := \{f \in \mathcal{A}_n : f/z \in \mathcal{P}_n\},$$

which was introduced by MacGregor [12].

Theorem 2.15. *The sharp $\mathcal{S}^*_{SG,n}$ radius for $\mathcal{W}_n$ is*

$$R_{\mathcal{S}^*_{SG,n}}(\mathcal{W}_n) := r_w = \left(\frac{e-1}{\sqrt{n^2(e+1)^2+(e-1)^2}+n(e+1)}\right)^{1/n}.$$

Proof. Let $f \in \mathcal{W}_n$, then there exists some $h \in \mathcal{P}_n$ such that $h(z) = f(z)/z$. Using Lemma 2.2, we have

$$\left|\frac{zh'(z)}{h(z)}\right| \leq \frac{2nr^n}{1-r^{2n}}.$$

This further implies that

$$\left|\frac{zf'(z)}{f(z)} - 1\right| = \left|\frac{zh'(z)}{h(z)}\right| \leq \frac{2nr^n}{1-r^{2n}}.$$

Now, by using 1.1 it follows that $f \in \mathcal{S}^*_{SG}$ provided $2nr^n/(1-r^{2n}) \leq (e-1)/(e+1)$. This inequality, after a few steps, reduces to the following

$$(e-1)r^{2n} + 2n(e+1)r^n - (e-1) \leq 0.$$

Solving this inequality gives us $r \leq r_w$. The sharpness of the result can be verified by the function $f_w(z) = z(1+z^n)/(1-z^n)$. Clearly this function satisfy $f(z)/z \in \mathcal{P}_n$ and thus $f_w \in \mathcal{W}_n$. Also, at the points $\pm r_w$ the function zf'_w/f_w touches the boundary of Δ_{SG}. ∎

For $\beta > 1$, the class $\mathcal{M}(\beta)$ introduced by Uralegaddi et al. [20] is given by

$$\mathcal{M}(\beta) := \left\{f \in \mathcal{A}_n : \operatorname{Re}\frac{zf'(z)}{f(z)} < \beta,\ z \in \mathbb{D}\right\}.$$

In terms of subordination, the above class can be written as

$$\mathcal{M}(\beta) := \left\{f \in \mathcal{A}_n : \frac{zf'(z)}{f(z)} \prec \frac{1+(1-2\beta)z^n}{1-z^n},\ z \in \mathbb{D}\right\}.$$

Theorem 2.16. *The $\mathcal{S}^*_{SG,n}$ radius for the class $\mathcal{M}(\beta)$ is given by*

$$\mathcal{S}^*_{SG}(\mathcal{M}_n(\beta)) = \left(\frac{e-1}{(e-1)+(e+1)(\beta)-1}\right)^{1/n}.$$

The result is sharp for the function $f(z) = z(1-z^n)^{(2(\beta-1)/n)}$.

The proof of the above theorem is omitted here as it is much similar to the proof of Theorem 2.14. Next we consider the class $\mathcal{C}(\alpha)$ $(0 \leq \alpha < 1)$, of convex functions of order α. Note that this class is a generalization of the class $\mathcal{C}$, which can be obtained by setting $\alpha = 0$.

Theorem 2.17. *Let $f \in \mathcal{S}^*_{SG}$, then f is convex of order α in $|z| < r_{\mathcal{C}}(\alpha)$, where $r_{\mathcal{C}}(\alpha) \in (0,1)$ is the smallest positive root of*

$$e^r(r+\alpha) - 2 + \alpha = 0.$$

Proof. Let $f \in \mathcal{S}^*_{SG}$, then $zf'(z)/f(z) \prec 2/(1+e^{-z})$ and thus there exists a function ω with $\omega(0) = 0$ and $|\omega(z)| < 1$ such that

$$\frac{zf'(z)}{f(z)} = \frac{2}{1+e^{-\omega(z)}}.$$

Differentiating the above equation logarithmically, we obtain

$$1+\frac{zf''(z)}{f'(z)}=\frac{2}{1+e^{-\omega(z)}}+\frac{\omega'(z)e^{-\omega(z)}}{1+e^{-\omega(z)}}.$$

Note that

$$\begin{aligned}\operatorname{Re}\left(1+\frac{zf''(z)}{f'(z)}\right) &= \operatorname{Re}\left(\frac{2}{1+e^{-\omega(z)}}+\frac{\omega'(z)e^{-\omega(z)}}{1+e^{-\omega(z)}}\right)\\ &\geq \operatorname{Re}\left(\frac{2}{1+e^{-\omega(z)}}\right)-\left|\frac{\omega'(z)e^{-\omega(z)}}{1+e^{-\omega(z)}}\right|\\ &\geq \frac{2}{1+e^{r}}-\frac{re^{r}}{1+e^{r}},\end{aligned}$$

which is greater than α provided $r\leq r_{\mathcal{C}}(\alpha)$. The result is sharp for the function

$$f_0(z)=z\exp\left(\int_0^z\frac{e^t-1}{t(e^t+1)}dt\right)=z+\frac{z^2}{2}+\frac{z^3}{8}+\frac{z^4}{144}-\frac{5z^5}{1152}+\cdots,$$

∎

Corollary 2.18. *The sharp radius of convexity for the functions in $\mathcal{S}^*_{SG}$ is $r\approx 0.852606$.*

References

[1] R. M. Ali, N. K. Jain and V. Ravichandran, *On the radius constants for classes of analytic functions*, Bull. Malays. Math. Sci. Soc. (2) **36** (2013), no. 1, 23–38.

[2] N. E. Cho, V. Kumar, S. S. Kumar and V. Ravichandran, *Radius problems for starlike functions associated with the sine function*, Bull. Iranian Math. Soc. **45** (2019), no. 1, 213–232.

[3] P. Goel and S. Sivaprasad Kumar,*Certain class of starlike functions associated with modified sigmoid function*, Bull. Malays. Math. Sci. Soc. **43** (2020), no. 1, 957–991.

[4] W. Janowski, *Extremal problems for a family of functions with positive real part and for some related families*, Ann. Polon. Math. **23** (1970/71), 159–177.

[5] R. Kargar, A. Ebadian and J. Sokół, *On Booth lemniscate and starlike functions*, Anal. Math. Phys. **9** (2019), no. 1, 143–154.

[6] M. G. Khan, B. Ahmad, G. Murugusundaramoorthy, R. Chinram, and W. K. Mashwani, *Applications of modified sigmoid functions to a class of starlike functions*, J. Funct. Spaces **2020**, doi.org/10.1155/2020/8844814

[7] K. Khatter, V. Ravichandran and S. Sivaprasad Kumar, *Starlike functions associated with exponential function and the lemniscate of Bernoulli*, Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM **113** (2019), no. 1, 233–253.

[8] S. S. Kumar and P. Goel, *Starlike functions and higher order differential subordinations*, Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM **114** (2020), no. 4, Paper No. 192, 23 pp.

[9] S. S. Kumar and G. Kamaljeet, *A cardioid domain and starlike functions*, Anal. Math. Phys. **11** (2021), no. 2, Paper No. 54, 34 pp.

[10] S. Kumar and V. Ravichandran, *A subclass of starlike functions associated with a rational function*, Southeast Asian Bull. Math. **40** (2016), no. 2, 199–212.

[11] W. C. Ma and D. Minda, *A unified treatment of some special classes of univalent functions*, in Proceedings of the Conference on Complex Analysis (Tianjin, 1992), 157–169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge, MA.

[12] T. H. MacGregor, *The radius of univalence of certain analytic functions*, Proc. Amer. Math. Soc. **14** (1963), 514–520.

[13] W. K. Mashwani, B. Ahmad, N. Khan, M. G. Khan, S. Arjika, B. Khan and R. Chinram, *Fourth Hankel determinant for a subclass of starlike functions based on modified sigmoid*, J. Funct. Spaces **2021**, doi.org/10.1155/2021/6116172

[14] R. Mendiratta, S. Nagpal and V. Ravichandran, *A subclass of starlike functions associated with left-half of the lemniscate of Bernoulli*, Internat. J. Math. **25** (2014), no. 9, 1450090, 17 pp.

[15] V. Ravichandran, F. Rønning and T. N. Shanmugam, *Radius of convexity and radius of starlikeness for some classes of analytic functions*, Complex Variables Theory Appl. **33** (1997), no. 1-4, 265–280.

[16] M. O. Reade, *On close-to-convex univalent functions*, Michigan Math. J. **3** (1955), 59–62.

[17] G. M. Shah, *On the univalence of some analytic functions*, Pacific J. Math. **43** (1972), 239–250.
[18] K. Sharma, N. K. Jain and V. Ravichandran, *Starlike functions associated with a cardioid*, Afr. Mat. **27** (2016), no. 5-6, 923–939.
[19] P. Sharma, R. K. Raina and J. Sokół, *Certain Ma-Minda type classes of analytic functions associated with the crescent-shaped region*, Anal. Math. Phys. **9** (2019), no. 4, 1887–1903.
[20] B. A. Uralegaddi, M. D. Ganigi and S. M. Sarangi, *Univalent functions with positive coefficients*, Tamkang J. Math. **25** (1994), no. 3, 225–230.
[21] L. A. Wani and A. Swaminathan, *Starlike and convex functions associated with a nephroid domain*, Bull. Malays. Math. Sci. Soc. **44** (2021), no. 1, 79–104.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: priyanka.goel0707@gmail.com

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

Noname manuscript No.
(will be inserted by the editor)

Radius Properties of Harmonic Mappings with Fixed Analytic Part

Kamaljeet Gangania

arXiv:2208.00757v1 [math.CV] 1 Aug 2022

Received: date / Accepted: date

Abstract We investigate harmonic mappings $f = h + \bar{g}$ defined in the unit disk, where g and h satisfy certain prescribed conditions and the analytic part h belongs to the Ma and Minda class of starlike functions. Certain sharp radius results for univalency, close-to-convexity, fully starlikeness, fully convexity, and radius of strongly starlikeness are established. We also calculate the radius of uniformly starlikeness and convexity for these functions. Several results enhance the well-known radius results.

Keywords Harmonic mappings · Ma-Minda starlike functions · Radius problems · Univalency · Close-to-convexity · Fully starlikeness and convexity · Uniformly starlike and convex functions

Mathematics Subject Classification (2010) 30C45 · 58E20 · 30A05 · 30C20 · 35C50

1 Introduction

A complex-valued function f is harmonic if and if $f \in C^2(\Omega)$ (continuous first and second partial derivatives in Ω) with $\Delta f = 4f_{z\bar{z}} = 0$. Suppose that f is a harmonic mapping of a proper simply connected domain $\Omega \subset \mathbb{C}$. Then $f : \Omega \to \mathbb{C}$ is a harmonic function if and only if the function f has the representation $f = h + \bar{g}$, where h and g are analytic in Ω. Therefore, well-known result of Lewy's [18] implies that such function f is locally univalent and sense-preserving in Ω if and only if $|h'(z)| > |g'(z)|$ in Ω. Equivalently, f

The work of Kamajeet Gangania is supported by University Grant Commission, New-Delhi, India under UGC-Ref. No.:1051/(CSIR-UGC NET JUNE 2017).

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
E-mail: gangania.m1991@gmail.com

is sense-preserving local homeomorphism if and only if $J_f(z) > 0$, where the jacobian $J_f(z) = |f_z(z)|^2 - |f_{\bar{z}}(z)|^2$. Further, by Riemann mapping theorem there is a conformal mapping φ of $\mathbb{D}$ onto Ω, and it follows that the composition $f \circ \varphi$ maps $\mathbb{D}$ harmonically onto Ω. As a consequence, we can assume that Ω is the unit disk and that f is sense-preserving in $\mathbb{D}$. Because $J_f > 0$ for sense-preserving harmonic mapping $f = h + \bar{g}$, $f_z(0) = h'(0) \neq 0$, and so we arrive at the normalized form $(f(z) - f(0))/f_z(0)$ of f. Thus, the analytic and co-analytic parts of the harmonic mapping $f = h + \bar{g}$ in $\mathbb{D}$ can be written, respectively,

$$h(z) = z + \sum_{n=2}^{\infty} a_n z^n \quad \text{and} \quad g(z) = \sum_{n=1}^{\infty} b_n z^n. \tag{1}$$

Denote by $\mathcal{S}_{\mathcal{H}}$ (see [5]) the class of all complex-valued harmonic univalent and sense-preserving mappings f in the unit disk $\mathbb{D}$ with $f(0) = 0 = f_z(0) = 1$. Let $\mathcal{A}$ denotes the class of analytic functions in $\mathbb{D}$ and normalized by the condition $h(0) = 0 = h'(0) - 1$. We note that $\mathcal{S}_{\mathcal{H}}$ reduces to $\mathcal{S}$, the class of normalized univalent analytic functions in $\mathbb{D}$ whenever the co-analytic part of f is zero. Furthermore, for $f = h + \bar{g} \in \mathcal{S}_{\mathcal{H}}$ with $g'(0) = b$ and $|b| < 1$ (because $J_f(0) = 1 - |b|^2 > 0$), the function $F = (f - \overline{bf})(1 - |b|^2) \in \mathcal{S}_{\mathcal{H}}$. The univalency of F is obtained by applying an affine mapping to f. Observe that $F_{\bar{z}}(0) = 0$. Therefore, we may restrict our attention to the subclass

$$\mathcal{S}^0_{\mathcal{H}} = \{f \in \mathcal{S}_{\mathcal{H}} : f_{\bar{z}}(0) = 0\}.$$

The condition $f_{\bar{z}}(0) = 0$ is equivalent to the assertion that the second complex dilatation $\omega_f(z) = \overline{f_z(z)}/f_z(z)$ of f is zero. Clearly, $\mathcal{S} \subset \mathcal{S}^0_{\mathcal{H}} \subset \mathcal{S}_{\mathcal{H}}$. Although both the classes $\mathcal{S}_{\mathcal{H}}$ and $\mathcal{S}^0_{\mathcal{H}}$ are known to be normal families, only $\mathcal{S}^0_H$ is compact with respect to the topology of locally uniform convergence. The geometric subclasses of $\mathcal{S}_{\mathcal{H}}$ consisting of convex, starlike and close-to-convex functions in $\mathbb{D}$ are denoted $\mathcal{C}_{\mathcal{H}}$, $\mathcal{S}^*_{\mathcal{H}}$ and $\mathcal{K}_{\mathcal{H}}$ respectively. Let $\mathcal{C}^0_{\mathcal{H}}$, $\mathcal{S}^{*0}_{\mathcal{H}}$ and $\mathcal{K}^0_{\mathcal{H}}$ denote the subclasses of $\mathcal{C}_{\mathcal{H}}$, $\mathcal{S}^*_{\mathcal{H}}$ and $\mathcal{K}_{\mathcal{H}}$ with the condition $f_{\bar{z}}(0) = 0$ respectively.

In 1992, Ma and Minda [19] introduced and studied the subclasses of $\mathcal{S}$ of starlike and convex univalent functions :

$$\mathcal{S}^*(\psi) = \left\{h \in \mathcal{A} : \frac{zh'(z)}{h(z)} \prec \psi(z)\right\} \text{ and } \mathcal{C}(\psi) = \left\{h \in \mathcal{A} : 1 + \frac{zh''(z)}{h'(z)} \prec \psi(z)\right\},$$

where the univalent function ψ defined as $\psi(z) = 1 + p_1 z + p_2 z^2 + \cdots$ has positive real in $\mathbb{D}$. The symbol $\prec$ denotes the usual subordination. Note that $\mathcal{S}^*((1+z)(1-z))$ and $\mathcal{C}((1+z)(1-z))$ respectively reduce to the well-known classes $\mathcal{S}^*$ and $\mathcal{C}$ of starlike and convex functions. For some current radius results, see [6,7,8,9,14,16,17,29] and their references.

For $f \in \mathcal{S}^*(\psi)$ or $\mathcal{C}(\psi)$, the starlikeness or convexity of the range set $f(\mathbb{D})$ is inherited by the level sets $f(\mathbb{D}_r)$ for all $0 < r < 1$. However, the hereditary property of starlikeness and convexity for harmonic function does not hold. For this, Chuaqui et al. [4] defined the notion of fully starlike and fully convex harmonic mappings.

Definition 1 [4] A harmonic function $f = h + \bar{g}$ with h and g of the form (1) is said to be fully starlike of order $0 \leq \beta < 1$ if each circle $C_r = \{z : |z| = r\}$ is one-to-one mapped onto a curve that bounds a starlike domain with respect to the zero and satisfy

$$\frac{\partial}{\partial\theta}(\arg f(re^{i\theta})) > \beta; \quad 0 \leq \theta < 2\pi,\ 0 < r < 1.$$

Similary, f is said to be fully convex of order β, if f maps one-to-one each circle C_r onto a curve that bounds a convex domain and satisfies

$$\frac{\partial}{\partial\theta}\{\arg(\frac{\partial}{\partial\theta}f(re^{i\theta}))\} > \beta; \quad 0 \leq \theta < 2\pi,\ 0 < r < 1.$$

In 1969, Bernardi generalized Alexander's theorem by introducing the function $\phi_\nu : \mathbb{D} \to \mathbb{C}$ defined by

$$\phi_\nu(z) = \frac{\nu+1}{z^\nu}\int_0^z \xi^{\nu-1}\phi(\xi)d\xi,$$

where ϕ is analytic in $\mathbb{D}$, with $\phi(0) = 0 = \phi'(0) - 1$. According to the Radó-Kaneser-Choquet theorem (see [5]), fully convex functions of order β are necessarily univalent in $\mathbb{D}$. However, fully harmonic starlike functions need not be univalent in $\mathbb{D}$. The well-known Alexander's relation $\phi \in \mathcal{S}^*$ if and only if $\int_0^z \phi(\xi)/\xi d\xi \in \mathcal{C}$ between the classes of starlike and convex function. Sheil-Small [28] extended this result to the harmonic mappings and proved that

Theorem 1 *[28](Sheil-Small) If $f = h + \bar{g} : \mathbb{D} \to \mathbb{C}$ fixes zero, is univalent, and has a starlike range, and H and G are the analytic functions in $\mathbb{D}$ defined by $zH'(z) = h(z)$, $zG'(z) = -g(z)$, $H(0) = G(0) = 0$, then $F = H + \overline{G}$ is univalent, and has a convex range.*

For $\nu \geq 0$, $\mu \in \mathbb{D}$ and $f \in \mathcal{S}_\mathcal{H}$, Muir [21] generalized the Bernardi integral operator to the harmonic functions, by defining $\bigwedge_{\nu,\mu}[f] : \mathbb{D} \to \mathcal{C}$ as follows:

$$\bigwedge_{\nu,\mu}[f](z) = \frac{\nu+1}{z^\nu}\int_0^z \xi^{\nu-1}h(\xi)d\xi + \mu\overline{\frac{\nu+1}{z^\nu}\int_0^z \xi^{\nu-1}g(\xi)d\xi}.$$

Clearly, $\bigwedge_{0,-1}$ is the harmonic analogue of Alexander's operator, given in Theorem 1. In fact, the classes $\mathcal{C}_\mathcal{H}$ and $\mathcal{S}^*_\mathcal{H}$ are preserved under the operator $\bigwedge_{0,-1}$. However, $\mathcal{C}_\mathcal{H}$ and $\mathcal{S}^*_\mathcal{H}$ are not necessarily preserved under $\bigwedge_{0,1}$. Indeed, $\bigwedge_{0,1}[K] \notin \mathcal{S}^{*0}_\mathcal{H}$, where the harmonic Koebe function $K(z)$ is given by

$$K(z) = \frac{z - \frac{1}{2}z^2 + \frac{1}{6}z^3}{(1-z)^3} + \overline{\frac{\frac{1}{2}z^2 + \frac{1}{6}z^3}{(1-z)^3}}.$$

There has been an interplay between the sense-preserving harmonic mappings and their analytic part. Mocanu [20] conjectured that if h and g are analytic functions in $\mathbb{D}$ that satisfy $g'(z) = zh'(z)$ and $\Re(1+zh''(z)/h'(z)) > -1/2$,

then the harmonic function $f = h+\bar{g}$ is univalent in $\mathbb{D}$. Bshouty and Lyssaik [3] proved that a sense-preserving harmonic function $f = h + \bar{g}$ where h and g of the form (1) is necessarily univalent in $\mathbb{D}$ if the dilatation of f is $w_f(z) = z$ and h satisfies $\Re(1 + zh''(z)/h'(z)) > -1/2$ for all $z \in \mathbb{D}$.

Recently, harmonic functions were studied with different dilatation and analytic part h belonging to different geometric subclasses, see [1,3,22,26,27]. Recall that a domain Ω is called close-to-convex if the complement of Ω can be written as a union of half-lines such that corresponding open half-lines are disjoint. A function $f \in \mathcal{S}_{\mathcal{H}}$ which maps $\mathbb{D}$ onto a close-to-convex domain is called a *close-to-convex function*. Clunie and Sheil-Small [5] proved that

Theorem 2 *[5] Let $f = h + \bar{g}$ be a locally univalent harmonic function in $\mathbb{D}$ and $h + \epsilon g$ $(|\epsilon| \leq 1)$ is convex in $\mathbb{D}$. Then f is univalent and close-to-convex in $\mathbb{D}$.*

Note that, in particular, if h is starlike in Theorem 2, then f need not be univalent in $\mathbb{D}$. The Koebe function $k(z) = z/(1-z)^2$ is starlike. However, the harmonic function given by

$$F(z) = \frac{z}{(1-z)^2} + \overline{\frac{\ln(1-z)}{2} + \frac{z(1-2z)}{2(1-z)^2}}$$

is sense-preserving in $\mathbb{D}$ as $|g'(z)/h'(z)| = |z(1+z)/2| < 1$ but not univalent in $\mathbb{D}$. Further, the harmonic function

$$G(z) = \frac{z}{(1-z)^2} + \overline{\frac{z}{2(1-z)}}$$

is sense-reversing in $\mathbb{D}$ as $|g'(-3/4)/h'(-3/4)| = 7/2 > 1$, but G is not univalent in $\mathbb{D}$ and also $g(z) = \frac{1}{2}(1-z)h(z)$. Hence, it is natural to determine the radius of univalency of harmonic functions which are not necessarily sense-preserving when $h \in \mathcal{S}^*(\psi)$ and $g(z) = \varphi(z)h(z)$, where $\varphi \in \mathcal{B}_s$, the class of analytic functions $\varphi : \mathbb{D} \to \mathbb{D}$ such that $|\phi(z)| \leq 1$ for all $z \in \mathbb{D}$.

We investigate the harmonic functions and the operator $\bigwedge_{0,1}$ for the radius of close-to-convexity and radius of univalency when the analytic part is a Ma-Minda starlike function (that is, $h \in \mathcal{S}^*(\psi)$) when the analytic and co-analytic part satisfy certain conditions. Further, we derive the hereditary radius of fully starlikeness, fully convexity, strongly starlikeness and strongly convexity of such harmonic functions. Unless otherwise stated, we assume that the function ψ has positive coefficients in its power series expansion throughout the discussion.

2 Harmonic mappings with Analytic part from the class of Ma-Minda starlike functions

The radius of univalency and the radius of close-to-convexity of complex-valued harmonic functions $f = h + \bar{g}$, where h and g are given by (1) satisfying the

coefficient conditions

$$|a_n| \leq \frac{(2n+1)(n+1)}{6} \quad \text{and} \quad |b_n| \leq \frac{(2n-1)(n-1)}{6}, \quad n \geq 2$$

is found to be 0.112903, see [15]. Recently, Ponnusamy and Kaliraj [22, Theorem 3.8] proved that for sense-preserving harmonic functions $f = h + \bar{g}$ such that h is univalent in $\mathbb{D}$, where h and g are given by (1), the radius of univalency and close-to-convexity equals to $2 - \sqrt{3} \approx 0.267949$.

Now, we obtain the radius of sense-preserving and univalency for the harmonic functions $f = h + \bar{g}$ such that $g(z) = \varphi(z)h(z)$ and $h \in \mathcal{S}^*(\psi)$, where $\varphi \in \mathcal{B}_s$. For this, we need to recall

Lemma 1 *[26] Let $r \in (0,1)$ and $f = h + \bar{g}$ be a sense-preserving harmonic mapping in $\mathbb{D}_r = \{z : |z| < r\}$. Further let h be univalent convex in $\mathbb{D}$. Then f is univalent in $\mathbb{D}_r$.*

Theorem 3 *Let $f = h + \bar{g}$ be the harmonic mapping in $\mathbb{D}$ such that $h \in \mathcal{S}^*(\psi)$ and $g(z) = \varphi(z)h(z)$, where $\varphi \in \mathcal{B}_s$. Further, assume that*

$$\min_{|z|=r} |\psi(z)| = m(r).$$

Then f is sense-preserving and univalent in $\mathbb{D}_R$, where $R = \min\{r_\psi, r_c\}$ and r_ψ is the unique postive root of the equation

$$(1 - r^2)m(r) - 2r = 0$$

and r_c is the radius of convexity of functions h in $\mathcal{S}^*(\psi)$. If $r_c \geq r_\psi$, then the result is sharp for the case $m(r) = \psi(-r)$.

Proof Since h is convex (univalent) in $|z| < r_c$. From Lemma 1, it suffices to prove that f is sense-preserving in $|z| < R$. As $h \in \mathcal{S}^*(\psi)$, by definition we have

$$\frac{zh'(z)}{h(z)} = \psi(\omega(z)), \quad \omega \in \mathcal{B}_s,$$

Since ψ is univalent, it follows by Maximum-minimum principle of modulus with the hypothesis $\min_{|z|=r} |\psi(z)| = m(r)$ that

$$\left| \frac{h(z)}{h'(z)} \right| \leq \frac{|z|}{\min_{|z|=r} |\psi(z)|} = \frac{r}{m(r)}. \tag{2}$$

Since $g(z) = \varphi(z)h(z)$ implies that

$$|g'(z)| = |h'(z)| \left\{ |\varphi(z)| + \left| \frac{h(z)}{h'(z)} \right| |\varphi'(z)| \right\}. \tag{3}$$

Now by employing the Schwarz-Pick inequality

$$|\varphi'(z)| \leq \frac{1 - |\varphi(z)|^2}{1 - |z|^2}$$

in the Eq. (3), and using (2) gives that $|g'(z)| \leq T(r)|h'(z)|$, where

$$T(r) := \left\{ |\varphi(z)| + \frac{r}{m(r)} \left(\frac{1-|\varphi(z)|^2}{1-|z|^2} \right) \right\}.$$

Now using [6, Theorem 5] we have $|T(r)| < 1$ such that

$$|g'(z)| < |h'(z)| \quad \text{in } |z| < r_\psi,$$

where the radius r_ψ is sharp and is the smallest root of the equation

$$(1-r^2)m(r) - 2r = 0$$

in $(0,1)$ This in view of Lemma 1, shows that f is sense-preserving and univalent in $|z| < R = \min\{r_c, r_\psi\}$. Further, if $r_c > r_\psi$ and $m(r) = \psi(-r)$, then [6, Proof of Theorem 2] implies that

$$|g'(z)| > |h'(z)| \quad \text{for all } z = r > r_\psi,$$

which completes the proof of sharpness. □

Now, for different choices of the function ψ, we expolre Theorem 3. Note that for harmonic functions $f = h + \bar{g}$ such that g is majorized by h and $h \in \mathcal{S}^*(\psi)$, we obtain the improved radius of univalency in Corollary 1. For sense-preserving harmonic functions with univalent analytic part, radius of univalency is $2-\sqrt{3} > 3-\sqrt{8}$ the conjectured radius of convexity of $f \in \mathcal{S}_H$, see [28]. Let us recall an known class of starlike functions defined by

$$\mathcal{S}^*[M] := \{h \in \mathcal{A} : \left| \frac{zf'(z)}{f(z)} - M \right| < M\}.$$

Such functions implies that $1 + \Re\frac{zh''(z)}{h'(z)} \geq \frac{(1+m)r}{(1-r)(1+mr)}$, where $m = 1 - \frac{1}{M}$.

Corollary 1 *Let $f = h+\bar{g}$ be the harmonic mapping in $\mathbb{D}$ such that $h \in \mathcal{S}^*(\psi)$ and $g(z) = \varphi(z)h(z)$, where $\varphi \in \mathcal{B}_s$. Then f is sense-preserving and univalent in $\mathbb{D}_R$, where $R = \min\{r_\psi, r_c\}$ and r_ψ is the unique postive root of the equation*

$$(1-r^2)m(r) - 2r = 0$$

and r_c is the radius of convexity of functions h in $\mathcal{S}^*(\psi)$. The result follows for each one of the following cases:

(*i*) $R = 2-\sqrt{3} = r_\psi = r_c \approx 0.267949$ *when* $\psi(z) = (1+z)/(1-z)$.
(*ii*) $R = r_\psi \approx 0.3524$ *is the root of* $(1-r)^3(1+r)^2 - 4r^2 = 0$ *for* $\psi(z) = \sqrt{1+z}$.
(*iii*) $R = r_\psi \approx 0.3237$ *is the root of* $(1-r^2) - 2re^r = 0$ *for* $\psi(z) = e^z$.
(*iv*) $R = r_\psi \approx 0.358473$ *is the root of* $(1-r^2)(\sqrt{1+r^2} - r) - 2r$ *when* $\psi(z) = z + \sqrt{1+z^2}$.
(*v*) $R = r_\psi \approx 0.358473$ *is the root of* $(1-r^2) - r(1+e^r) = 0$ *for* $\psi(z) = 2/(1+e^{-z})$.
(*vi*) $R = r_c \approx 0.345$ *for* $\psi(z) = 1 + \sin z$. *(see, [17, Corollary 2.4])*
(*vii*) $R = r_\psi$ *is the root of* $(1+r)(1-r)^2 - 2r(1+mr) = 0$ *when* $h \in \mathcal{S}^*[M]$.

Next to study the radius of fully starlikeness and convexity for functions in $\mathcal{S}^0_{\mathcal{H}}$ with the condition $g(z) = \varphi(z)h(z)$, we invoke a result of Jahangiri.

Lemma 2 *[12,13] Let the harmonic function f be given by $f(z) = h(z) + \overline{g(z)} = z + \sum_{n=2}^{\infty} a_n z^n + \overline{\sum_{n=2}^{\infty} b_n z^n}$ in $\mathbb{D}$ and $0 \leq \beta < 1$.*

(i) The function f is fully starlike of order β, if

$$\sum_{n=2}^{\infty}(n-\beta)|a_n| + \sum_{n=2}^{\infty}(n+\beta)|b_n| \leq 1-\beta.$$

(ii) The function f is fully convex of order β, if

$$\sum_{n=2}^{\infty} n(n-\beta)|a_n| + \sum_{n=2}^{\infty} n(n+\beta)|b_n| \leq 1-\beta.$$

Remark 1 In the case when the coefficient of power series of the function ψ in the class $\mathcal{S}^*(\psi)$ are negative, then without loss of generality, we can replace the function $h_\psi(z) = z\exp\int_0^z(\psi(t)-1)/tdt := z + \sum_{n=2}^{\infty} a_n z^n$ by $\hat{h}_\psi(z) := z + \sum_{n=2}^{\infty} |a_n| z^n$ and same for its derivative, in what follows.

Theorem 4 *Let $f = h + \bar{g}$ be a harmonic function such that $h \in \mathcal{S}^*(\psi)$ in $\mathbb{D}$ and $g(z) = \varphi(z)h(z)$, where $\varphi \in \mathcal{B}_s$, h and g are given by (1). Then for $\beta \in [0,1)$, f is fully starlike of order β in $|z| < R_\beta = \min\{1/3, r_\psi\}$, where r_ψ is the smallest postive root of the equation*

$$4h'_\psi(r) + \beta = 5$$

in $(0,1)$, where $h_\psi(z) = z\exp\left(\int_0^z(\psi(t)-1)/tdt\right)$.

Proof Let $0 < r < 1$. Consider the function

$$f_r(z) = \frac{1}{r}f(rz) = z + \sum_{n=2}^{\infty} a_n r^{n-1} z^n + \overline{\sum_{n=2}^{\infty} b_n r^{n-1} z^n}.$$

Now, it suffices to prove that the function f_r is fully starlike of order β in $|z| \leq R_\beta$. Since $g(z) = \varphi(z)h(z)$, where $\varphi \in \mathcal{B}_s$, it follows from [8, Corollary 1] that in $|z| \leq 1/3$

$$|b_n| \leq |a_n|. \tag{4}$$

Invoking the result [8, Corollary 1] for the subordination $h(z)/z \prec h_\psi(z)/z$ (see [19]) gives together with inequality (4) that for the disk $|z| \leq 1/3$

$$|a_n| \leq |\hat{a_n}|, \tag{5}$$

where $h_\psi(z) = z\exp\left(\int_0^z(\psi(t)-1)/t dt\right) = z+\sum_{n=2}^{\infty}\hat{a_n}z^n$. Therefore, from (4) and (5)

$$\begin{aligned} S_1 &= \sum_{n=2}^{\infty}\left(\frac{n-\beta}{1-\beta}\right)|a_n|r^{n-1} + \sum_{n=2}^{\infty}\left(\frac{n-\beta}{1-\beta}\right)|b_n|r^{n-1} \\ &\leq \sum_{n=2}^{\infty}\left(\frac{n-\beta}{1-\beta}\right)|\hat{a_n}|r^{n-1} + \sum_{n=2}^{\infty}\left(\frac{n-\beta}{1-\beta}\right)|\hat{a_n}|r^{n-1} \\ &= \frac{4}{1-\beta}\sum_{n=2}^{\infty}n|\hat{a_n}|r^{n-1} = \frac{4}{1-\beta}(h'_\psi(r)-1) \end{aligned}$$

holds in $|z| \leq 1/3$. Hence, by Lemma 2, we see that $S_1 \leq 1$ holds in $|z| \leq R_\beta = \min\{1/3, r_\psi\}$, where r_ψ is the smallest positive root of the equation

$$\frac{4}{1-\beta}(h'_\psi(r)-1) = 1,$$

which implies that f is fully starlike of order β in $|z| \leq R_\beta$. □

Corollary 2 *Let f satisfies the hypothesis of Theorem 4. Then for $\beta \in [0,1)$, f is fully starlike of order β in $|z| < R_\beta = \min\{1/3, r_\psi\}$, where r_ψ is the smallest postive root of the equation:*

(i) $4h'_\psi(r) + \beta = 5$ *for* $\psi(z) = \frac{1+Dz}{1+Ez}$, *where* $-1 \leq E < D \leq 1$ *and*

$$h'_\psi(r) = \begin{cases} \psi(r)(1+Er)^{\frac{D-E}{E}}, & E \neq 0; \\ \psi(r)e^{Dr}, & E = 0. \end{cases}$$

(ii) $4\psi(r)e^{(\int_0^r \frac{\psi(t)-1}{t}dt)} + (\beta-5) = 0$ *for* $\psi(z) = (\frac{1+z}{1-z})^\eta$, *where* $0 < \eta \leq 1$.

(iii) $(16\sqrt{1+r})e^{2(\sqrt{1+r}-1)} + (\beta-5)(\sqrt{1+r}+1)^2 = 0$ *for* $\psi(z) = \sqrt{1+z}$ *and* $r_\psi < 1/3$ *for all* β.

(iv) $4(1+re^r)e^{e^r-1} + (\beta-5) = 0$ *for* $\psi(z) = 1+ze^z$ *and* $r_\psi < 1/3$ *for all* β.

(v) $8\psi(r)e^{(\psi(r)-1)} + (\beta-5)(\sqrt{1+r^2}+1) = 0$ *for* $\psi(z) = z+\sqrt{1+z^2}$ *and* $r_\psi < 1/3$ *for all* β.

Corollary 3 *Let f satisfies the hypothesis of Theorem 4. Then for $\beta \in [0,1)$, F given by*

$$F(z) = \bigwedge_{0,1}[f](z) = \int_0^z \frac{h(\xi)}{\xi}d\xi + \overline{\int_0^z \frac{g(\xi)}{\xi}d\xi}$$

is fully starlike of order β in $|z| < R_\beta = \min\{1/3, r_\psi\}$, where r_ψ is the smallest postive root of the equation

$$\frac{2}{1-\beta}\left(\frac{h_\psi(r)-r}{r}\right) = 1$$

in $(0,1)$, where $h_\psi(z) = z\exp\left(\int_0^z(\psi(t)-1)/t dt\right)$. See, Figure 1.

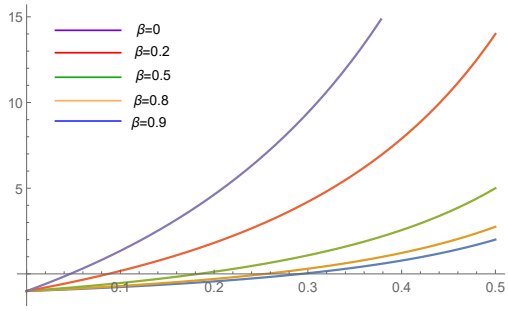


Fig. 1 Graph for the roots of equation for various choices of β with $\psi(z) = (1+z)/(1-z)$.

Hernandez and Martin [11] (also see, [25, Theorem 2.7]) proved that a sense-preserving harmonic function $f = h + \bar{g}$ is fully starlike in $\mathbb{D}$ if the analytic functions $h + \epsilon g$ are starlike in $\mathbb{D}$ for each $|\epsilon| = 1$. Now combining this result with a majorization result [6, Theorem 2] of Gangania and Kumar, we get that

Theorem 5 *Let h and g are given by (1) and $f = h + \bar{g}$ be a harmonic function such that $h \in \mathcal{S}^*(\psi)$ in $\mathbb{D}$ and $g(z) = \varphi(z)h(z)$, where $\varphi \in \mathcal{B}_s$. Further let $m(r) = \min_{|z|=r} \Re\psi(z)$. Then for $\beta \in [0,1)$, f is fully starlike of order β in $|z| < R_\beta = \min\{r_1, r_2\}$, where r_1 and r_2, respectively are the smallest positive roots of the following equations*

$$m(r) - \frac{r}{1-r} = \beta$$

and

$$(1-r^2)m(r) - 2r = 0$$

in $(0,1)$, where $h_\psi(z) = z\exp\left(\int_0^z (\psi(t)-1)/t dt\right)$.

In Figure 2, the left side graph is for $|z| < 0.238778$ in View of Theorem 4 and right side is for $|z| < 0.3524$ in view of Theorem 5 for the case when $\psi(z) = \sqrt{1+z}$.

In Theorem 5 if we include the condition that f is sense-preserving harmonic function then the radius can be further improved, which is stated in the following result.

Corollary 4 *Let $f = h + \bar{g}$ be a sense-preserving harmonic function such that $h \in \mathcal{S}^*(\psi)$ in $\mathbb{D}$. Further let $m(r) = \min_{|z|=r} \Re\psi(z)$. Then for $\beta \in [0,1)$, f is fully starlike of order β in $|z| < R_\beta$, which is the smallest positive root of the following equation*

$$m(r) - \frac{r}{1-r} = \beta.$$

For the choice of $\psi(z) = \sqrt{1+z}$ in above corollary, gives $R_0 = 0.43016 > r_2 = 0.3524$, where r_2 is defined in Theorem 5.

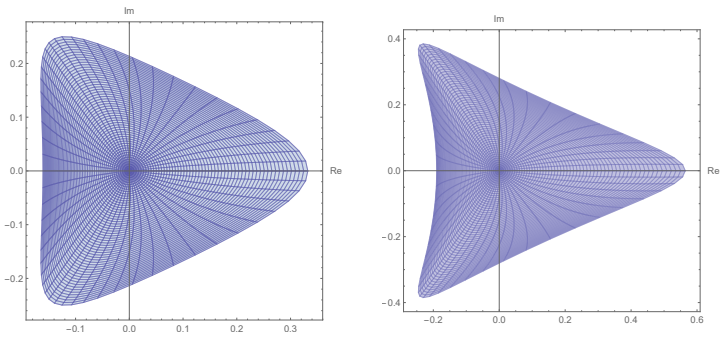


Fig. 2 Images of Harmonic function $f_0(z)+\overline{zf_0(z)}$ with $f_0(z)=\frac{4z\exp(2\sqrt{1+z}-2)}{(1+\sqrt{1+z})^2}$.

Theorem 6 *Let $f=h+\bar{g}$ be a harmonic function such that $h\in\mathcal{S}^*(\psi)$ in $\mathbb{D}$ and $g(z)=\varphi(z)h(z)$, where $\varphi\in\mathcal{B}_s$, h and g are given by (1). Then for $\beta\in[0,1)$, f is fully convex of order β in $|z|<C_\beta=\{1/3,r_\psi\}$, where r_ψ is the smallest postive root of the equation*

$$G'(r)*h'_\psi(r)=\frac{5-\beta}{4}$$

in $(0,1)$, where $G(z)=z/(1-z)$ and the symbol "$*$" represents the convolution.

Proof Similar to the proof of Theorem 4, it is enough to show that f_r is fully convex in $|z|\le C_\beta$. Now using the result [8, Corollary 1] on $g(z)=\varphi(z)h(z)$ and (5), we see that

$$\begin{aligned}S_2&=\sum_{n=2}^{\infty}\left(\frac{n(n-\beta)}{1-\beta}\right)|a_n|r^{n-1}+\sum_{n=2}^{\infty}\left(\frac{n(n+\beta)}{1-\beta}\right)|b_n|r^{n-1}\\&\le\sum_{n=2}^{\infty}\left(\frac{n(n-\beta)}{1-\beta}\right)|\hat{a_n}|r^{n-1}+\sum_{n=2}^{\infty}\left(\frac{n(n+\beta)}{1-\beta}\right)|\hat{a_n}|r^{n-1}\\&=\frac{4}{1-\beta}\sum_{n=2}^{\infty}n^2|\hat{a_n}|r^{n-1}\end{aligned}$$

holds in $|z|\le 1/3$. Note that using convolution, we can re-write

$$\sum_{n=2}^{\infty}n^2|\hat{a_n}|r^{n-1}=(G'(r)-1)*(h'_\psi(r)-1),$$

which implies that $S_2\le 1$ holds in $|z|\le C_\beta=\min\{1/3,r_\psi\}$, where r_ψ is the least positive root in $(0,1)$ of the equation

$$(G'(r)-1)*(h'_\psi(r)-1)=\frac{1-\beta}{4},$$

where $G(z)=z/(1-z)$. Hence, Lemma 2 says that f is fully convex of order β in $|z|\le C_\beta$. □

Corollary 5 *Let f satisfies the hypothesis of Theorem 6. Then for $\beta \in [0,1)$, F given by*

$$F(z) = \bigwedge_{0,1}[f](z) = \int_0^z \frac{h(\xi)}{\xi} d\xi + \overline{\int_0^z \frac{g(\xi)}{\xi} d\xi}$$

is fully convex of order β in $|z| < C_\beta = \min\{1/3, r_\psi\}$, where r_ψ is the smallest postive root of the equation

$$2h'_\psi(r) + \beta = 3$$

in $(0,1)$, where $h_\psi(z) = z\exp\left(\int_0^z (\psi(t)-1)/t dt\right)$.

Corollary 6 *Let f satisfies the conditions of Theorem 6. Then for $\beta \in [0,1)$, f is fully convex of order β in $|z| < C_\beta = \{1/3, r_\psi\}$, where r_ψ is the smallest positive root of the equation*

$$G'(r) * h'_\psi(r) = \frac{5-\beta}{4}$$

in $(0,1)$, where $G(z) = z/(1-z)$ and

(i) for $\psi(z) = \frac{1+Dz}{1+Ez}$, where $-1 \le E < D \le 1$:

$$h'_\psi(r) = \begin{cases} \psi(r)(1+Er)^{\frac{D-E}{E}}, & E \neq 0; \\ \psi(r)e^{Dr}, & E = 0. \end{cases}$$

(ii) for $\psi(z) = \sqrt{1+z}$: $h'_\psi(r) = \frac{4\sqrt{1+r}}{(\sqrt{1+r}+1)^2} e^{2(\sqrt{1+r}-1)}$.

(iii) for $\psi(z) = 1 + ze^z$: $h'_\psi(r) = (1+re^r)e^{e^r-1}$.

(iv) for $\psi(z) = z + \sqrt{1+z^2}$: $h'_\psi(r) = \frac{2(r+\sqrt{1+r^2})}{\sqrt{1+r^2}+1} e^{(r+\sqrt{1+r^2}-1)}$.

The following table show that the radius $C_\beta = r_\psi < 1/3$ in part *(iii)* of Corollary 6 is sharp, see Figure 3.

	S_5	S_{10}	S_{20}	$S_{n\to\infty}$
$\beta = 0$	0.1952	0.195106	0.195106	0.195106
$\beta = 1/2$	0.181806	0.181742	0.181742	0.181742
$\beta = 0.9$	0.17048	0.170435	0.170435	0.170435
				$r_\psi < 1/3$

Table 1 Behaviour of the roots of the equation $G'(r) * h'_\psi(r) = \frac{5-\beta}{4}$ when $\psi(z) = 1 + ze^z$ in Corollary 6.

For the sense-preserving harmonic functions $f = h + \bar{g}$ such that $f_{\bar{z}}(0) = 0$ and $h \in \mathcal{K}(c)$, where

$$\mathcal{K}(c) := \left\{ 1 + \frac{zh''(z)}{h'(z)} > c, \ -\frac{1}{2} \le c < 1 \right\}$$

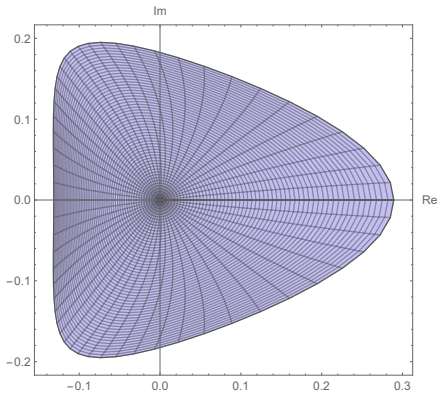


Fig. 3 Image of the harmonic function $z\exp(e^z-1)+\overline{z^2\exp(e^z-1)}$ in $|z|\leq 0.195106$.

is the class of convex functions of order c (such functions are univalent and convex in one direction in $\mathbb{D}$), Ponnusamy and Kaliraj [22, Theorem 4.6] proved that f is univalent and close-to-convex in $|z|<r_c$, where $r_c=1/3$ for $-1/2\leq c\leq c_0\approx -0.391827$, and $r_c=\cos\pi c$ for $c_0\leq c\leq 0$, where c_0 is the root of $\cos\pi c=1/3$.

Let us now consider the function $\psi(z)=\sqrt{1+z}$. Then $\mathcal{S}^*(\psi)$ reduces to the class of lemniscate starlike functions, denoted by $\mathcal{SL}^*$.

Example 1 Let $h\in\mathcal{SL}^*$ and $\varphi\in\mathcal{B}_s$. If $g'(z)=\varphi(z)h'(z)$ for all $z\in\mathbb{D}$, then the harmonic mappings $f=h+\bar{g}$ are univalent and close-to-convex in $|z|<r_0\approx 0.3119$ is the unique positive roof of $1-6r+12r^2-11r^3+4r^4=0$ in $(0,1)$.

Proof Let $Q_\delta(z)=h(z)-\delta g(z)$, where $|\delta|\leq 1$. Then we have

$$Q'_\delta(z)=(1-\delta\varphi(z))h'(z),$$

which further gives

$$1+\frac{zQ''_\delta(z)}{Q'_\delta(z)}=1+\frac{zh''(z)}{h'(z)}-\frac{\delta z\varphi'(z)}{1-\delta\varphi(z)}. \tag{6}$$

Let $p(z)=zh'(z)/h(z)$, then using Lemma [29], we get the sharp inequality

$$\Re\left(1+\frac{zh''(z)}{h'(z)}\right)\geq\sqrt{1-r}-\frac{r}{\sqrt{1-r}}. \tag{7}$$

Using inequality (7) in (6), we get

$$\Re\left(1+\frac{zQ''_\delta(z)}{Q'_\delta(z)}\right)\geq\sqrt{1-r}-\frac{r}{\sqrt{1-r}}-\Re\left(\frac{\delta z\varphi'(z)}{1-\delta\varphi(z)}\right). \tag{8}$$

The Schwarz-Pick inequality says that $\varphi \in \mathcal{B}_s$ satisfies

$$|\varphi'(z)| \leq \frac{1-|\varphi(z)|^2}{1-|z|^2}, \quad \text{which gives} \quad \left|\frac{\delta\varphi'(z)}{1-\delta\varphi(z)}\right| \leq \frac{|z|}{1-|z|}.$$

Hence,

$$\Re\left(\frac{\delta\varphi'(z)}{1-\delta\varphi(z)}\right) \leq \sqrt{1-r} - \frac{r}{\sqrt{1-r}}$$

for $|z| = r < r_0$, where $r_0 \approx 0.3119$ is the smallest positive roof of $1 - 6r + 12r^2 - 11r^3 + 4r^4 = 0$. Hence, in view of (8), f is univalent and close-to-convex in $|z| < r_0$. □

If the dilatation $\varphi = g'/h'$ of harmonic mapping $f = h + \bar{g}$ has the form $\varphi(z) = e^{i\theta}z^n$ $(\theta \in \mathbb{R}), n \in \mathbb{N}$, then we get

Example 2 Let $f = h + \bar{g}$ be the harmonic mapping in $\mathbb{D}$ such that $h \in \mathcal{SL}^*$ and $g'(z) = e^{i\theta}z^n h'(z)$. Then f is univalent and close-to-convex in $|z| < r_n$, where r_n is the smallest positive root of the equations

$$(1 - 4r + 4r^2) - r^n(2 - 8r + 8r^2) + r^{2n}(1 - n^2 + (n^2 - 4)r + 4r^2) = 0$$

in $(0, 1)$.

We now generalize Example 1. In fact, the following result can also be compared with [22, Theorem 3.8]. For certain results for the harmonic function $f = h + \bar{g}$ to be univalent and close-to-convex, we refer to see [1,2,3].

Theorem 7 *Let $h \in \mathcal{S}^*(\psi)$ and $\varphi \in \mathcal{B}_s$. If $g'(z) = \varphi(z)h'(z)$ for all $z \in \mathbb{D}$, then the harmonic functions $f = h + \bar{g}$ are univalent and close-to-convex in $|z| < r = \min\{r_0, r_c\}$, where r_c is radius of convexity for the class $\mathcal{S}^*(\psi)$ and r_0 is the minimal positive root of the equation*

$$(1 - r)H_\psi(r) - r = 0$$

with

$$H_\psi(r) := \min_{0<|z|=r} \Re\left(1 + \frac{zh''(z)}{h'(z)}\right).$$

Proof In Theorem 1, we notice that one requires the sharp inequality (7) in case when $h \in \mathcal{S}^*(\psi)$ which deals with the finding of radius of convexity for functions h. So without loss of generality, we may assume a continuous function $H_\psi(r)$ such that there exists some $r_c \in (0, 1)$ in such a way that $H_\psi(r) \geq 0$ for $r \in [0, r_c]$ and $H_\psi(r) < 0$ for $r \in (r_c, r_c + \epsilon)$ for some $\epsilon > 0$, and the following sharp inequality holds

$$\Re\left(1 + \frac{zh''(z)}{h'(z)}\right) \geq H_\psi(r) > 0, \quad \text{for all} \quad |z| = r \in [0, r_c),$$

where $H_\psi(0) = 1$. Hence, further following the footsteps of the Proof of Example 1, we get the desired result. □

Corollary 7 *Let $h \in \mathcal{S}^*_{SG}$ in Theorem 7. If $g'(z) = \varphi(z)h'(z)$ for all $z \in \mathbb{D}$, where $\varphi \in \mathcal{B}_s$ then the harmonic mappings $f = h + \bar{g}$ are univalent and close-to-convex in $|z| < r_0 \approx 0.3729$, where r_0 is the unique positive root of the equation*

$$\frac{2 - re^r}{1 + e^r} - \frac{r}{1 - r} = 0.$$

Let us now recall a basic definition to proceed further.

Definition 2 A locally univalent function $f = h + \bar{g}$ is said to be uniformly starlike in $\mathbb{D}$, if f is fully starlike in $\mathbb{D}$, and maps every circular arc Γ_ζ contained in $\mathbb{D}$ with center $\zeta \in \mathbb{D}$ onto the arc $f(\Gamma_\zeta)$ which is starlike with respect to ζ. Similarly, we have definition of uniformly convex function $f = h + \bar{g}$ in $\mathbb{D}$.

Recall the following important sufficient conditions for functions to be uniformly starlike and convex:

Lemma 3 *[23] Let the harmonic function $f = h + \bar{g} \in \mathcal{S}^0_{\mathcal{H}}$. Then the following holds:*

(i) The function f is uniformly starlike, if $\sum_{n=2}^{\infty} n(|a_n| + |b_n|) \leq 1/2$.
(ii) The function f is uniformly convex, if $\sum_{n=2}^{\infty} n(2n-1)(|a_n| + |b_n|) \leq 1$.

Using the Lemma 3, for harmonic functions $f \in \mathcal{S}^0_{\mathcal{H}}$, radius problems for the operator $\bigwedge_{0,1}[f]$ were studied in [10] with the prescribed conditions on the coefficient bounds $|a_n|$ and $|b_n|$ for the analytic and co-analytic part. We now prove the result when g is majorized by some function h in $\mathcal{S}^*(\psi)$.

Theorem 8 *Let $f = h + \bar{g} \in \mathcal{S}^0_{\mathcal{H}}$ such that $g(z) = \varphi(z)h(z)$ and $h \in \mathcal{S}^*(\psi)$ in $\mathbb{D}$, where $\varphi \in \mathcal{B}_s$. Then f is uniformly convex on the disk $|z| < r_{uc} = \min\{1/3, r_\psi\}$, where r_ψ is the smallest positive root of the equation*

$$2(G'(r) - 1) * (h'_\psi(r) - 1) - (h'_\psi(r) - 1) = \frac{1}{2}$$

in $(0, 1)$, where $G(z) = z/(1 - z)$ and $h_\psi(z) = z \exp\left(\int_0^z (\psi(t) - 1)/t dt\right)$.

Proof For $0 < r < 1$, let

$$f_r(z) = r^{-1}(rz) = z + \sum_{n=2}^{\infty} a_n r^{n-1} z^n + \overline{\sum_{n=2}^{\infty} b_n r^{n-1} z^n}, \quad (z \in \mathbb{D}).$$

In view of Lemma 3, let us consider the summation

$$S = \sum_{n=2}^{\infty} |a_n| r^{n-1} + \sum_{n=2}^{\infty} |b_n| r^{n-1}.$$

Thus, from (4) and (5), it follows that

$$S \leq \sum_{n=2}^{\infty} 2n(2n-1)|\hat{a_n}| r^{n-1} =: 2[2(G'(r) - 1) * (h'_\psi(r) - 1) - (h'_\psi(r) - 1)]$$
$$\leq 1$$

holds in $|z| = r \leq \min\{1/3, r_\psi\}$, where r_ψ is the least positive root of the equation

$$2(G'(r) - 1) * (h'_\psi(r) - 1) - (h'_\psi(r) - 1) = \frac{1}{2}.$$

This proves that f is uniformly convex on the disk $|z| < r_{uc}$. □

Theorem 9 *Let f satisfies the hypothesis of Theorem 8. Then f is uniformly starlike on the disk $|z| < r_{us} = \min\{1/3, r_\psi\}$, where r_ψ is the smallest positive root of the equation*

$$h'_\psi(r) - \frac{5}{4} = 0$$

in $(0, 1)$, where $h_\psi(z) = z \exp\left(\int_0^z (\psi(t) - 1)/t dt\right)$.

Proof It follows using the technique of Theorem 8 with inequalities (4), (5) and the application of Lemma 3. □

Corollary 8 *Let f satisfies the hypothesis of Theorem 8. Then F given by*

$$F(z) = \bigwedge_{0,1}[f](z) = \int_0^z \frac{h(\xi)}{\xi} d\xi + \overline{\int_0^z \frac{g(\xi)}{\xi} d\xi}$$

is uniformly starlike on the disk $|z| < r_{us} = \min\{1/3, r_\psi\}$, where r_ψ is the smallest positive root of the equation

$$4h_\psi(r) - 5r = 0$$

in $(0, 1)$, where $h_\psi(z) = z \exp\left(\int_0^z (\psi(t) - 1)/t dt\right)$.

The above corollary yields that

Corollary 9 *Let f satisfies the hypothesis of Theorem 8 with $\psi(z) = 1 + ze^z$. Then F given by*

$$F(z) = \bigwedge_{0,1}[f](z) = \int_0^z \frac{h(\xi)}{\xi} d\xi + \overline{\int_0^z \frac{g(\xi)}{\xi} d\xi}$$

is uniformly starlike on the disk $|z| < r_{us} \approx 0.201424$ is the unique positive root of the equation

$$4r \exp(e^r - 1) - 5r = 0.$$

In 2022, the notion of hereditarily strongly starlike harmonic functions appeared.

Definition 3 [24] Let f be the complex-valued harmonic function $f = h + \bar{g}$ with h and g of the form (1). The function f is called it hereditarly strongly starlike of order $\alpha \in (0, 1)$ if it is sense-preserving and univalent on $\mathbb{D}$ and if $f(\mathbb{D}_r)$ is a strongly starlike domain of order α for each $0 < r < 1$.

For such functions sharp bounds for $|a_n|$ and $|b_n|$ are yet to be known. We determine the radius of hereditarliy strongly starlikeness of order α.

Theorem 10 *Let the harmonic functions $f = h + \bar{g} \in \mathcal{S}^0_{\mathcal{H}}$ such that $g(z) = \varphi(z)h(z)$ and $h \in \mathcal{S}^*(\psi)$ in $\mathbb{D}$, where $\varphi \in \mathcal{B}_s$. Then $f \in \mathcal{SS}^0_{\mathcal{H}}(\alpha)$ on the disk $|z| < r_{ss} = \min\{1/3, r_\psi(\alpha)\}$, where $r_\psi(\alpha)$ is the smallest positive root of the equation*

$$\left(\frac{h_\psi(r) - r}{r}\right) * M_\alpha(r) - 2\sin\frac{\pi\alpha}{2} = 0,$$

where $0 < \alpha < 1$, and $M_\alpha(r) = \sum_{n=2}^{\infty}(A_n(\alpha) + B_n(\alpha))r^{n-1}$ with $A_n(\alpha) = n - 1 + |n - e^{-i\pi\alpha}|$ and $B_n(\alpha) = n + 1 + |n + e^{i\pi\alpha}|$.

Proof The result follows by invoking [24, Theorem 4.2], and proceeding similar to the proof of Theorem 8. □

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. S. V. Bharanedhar and S. Ponnusamy, Coefficient conditions for harmonic univalent mappings and hypergeomteric mappings, Rocky Mt. J. Math. (2014)
2. D. Bshouty, S. S. Joshi and S. B. Joshi, On close-to-convex harmonic mappings, Complex Var. Elliptic Equ. 58(4), 1195–1199 (2013)
3. D. Bshouty and A. Lyzzaik, Close-to-convexity criteria for planer harmonic mappings. Complex Anal. Oper. Theory, 5, 767–774 (2011)
4. M. Chuaqui, P. Duren and B. Osgood, Schwarzian derivatives and uniform local univalence, Comput. Methods Funct. Theory, 8, 21–34 (2008)
5. J. Clunie and T. Sheil-Small, Harmonic univalent functions, Ann. Acad. Sci. Fenn. Ser. A I Math, 9, 3–25, (1984)
6. K. Gangania and S. S. Kumar, On Certain Generalizations of $\mathcal{S}^*(\psi)$, Comput. Methods Funct. Theory (2021), https://doi.org/10.1007/s40315-021-00386-5.
7. K. Gangania and S. S. Kumar, Bohr radius for some classes of Harmonic mappings, Iran J Sci Technol Trans Sci (2022), https://doi.org/10.1007/s40995-022-01304-7
8. K. Gangania and S.S. Kumar, Bohr-Rogosinski phenomenon for $\mathcal{S}^*(\psi)$ and $\mathcal{C}(\psi)$, Accepted in Mediterranean Journal of Mathematics, (2021).
9. K. Gangania, Bohr Phenomenon for K-Quasiconformal harmonic mappings and Logarithmic Power Series (2021), https://arxiv.org/abs/2110.13068
10. N. Gosh and A. Vasudevarao, The radii of fully starlikeness and fully convexity of a harmonic operator, Monatsh Math, 188, 653-666 (2019)
11. R. Hernández and M. J. Martín, Stable geometric properties of analytic and harmonic functions, Preprint. Available from http://www.uam.es/mariaj.martin.
12. J. M Jahangiri, Coefficient bounds and univalence criterion for harmonic functions with negative coefficients, Ann Univ Mariae Cruie-Sklodowska Sec A, 52, 57–66
13. J. M Jahangiri, Harmonic starlike in the unit disk, J Math Anal Appl, **235** (2), 470–477 (1999)
14. Janowski, W.: Extremal problems for a family of functions with positive real part and for some related families. Ann. Polon. Math. 23, 159–177(1970)
15. D. Kalaj, S. Ponnusamy and M. Vuorinen, Radius of close-to-convexty and fully starlikeness of harmonic mappings, Complex Var. Elliptic Equ. 59(4), 539–552 (2014)
16. S. S. Kumar and G. Kamaljeet, A cardioid domain and starlike functions, Anal. Math. Phys. **11** (2021), no. 2, 54.

17. Cho, N. E., Kumar, V., Kumar, S. S., Ravichandran, V.: Radius problems for starlike functions associated with the sine function. Bull. Iranian Math. Soc. 45(1), 213–232(2019)
18. H. Lewy, On the nonvanishing of the Jacobian in certain one-to-one mappings, Bull. Am. Math. Soc. 42, 689-692 (1936)
19. W. C. Ma and D. Minda, A unified treatment of some special classes of univalent functions. Proceedings of the Conference on Complex Analysis, Tianjin, Conf Proc Lecture Notes Anal, I Int Press, Cambridge, MA, 157–169 (1992)
20. P. T. Mocanu, Injectivity conditions in the complex plane, Complex Anal. Oper. Theory, 5(3), 759–766 (2011)
21. S. Muir, Convex combinations of planar harmonic mappings realize through convolutions with half-stri mappings, Bull. Malays. Math. Sci. Soc., 40(2), 857–880 (2017)
22. S. Ponnusamy and A. S. Kaliraj, Constants and Characterization for Certain Classes of Univalent Harmonic Mappings, Mediterr. J. Math., 12 647-665 (2015)
23. S. Ponnusamy, J. K. Prajapat and S. Sairam, Uniformly starlike and uniformly convex harmonic mappings, J. Anal., 23, 121–129 (2015)
24. X. S. Ma, S. Ponnusamy and T. Sugawa, Harmonic spirallike functions and harmonic strongly starlike functions, Monatshefte för Mathematik, https://doi.org/10.1007/s00605-022-01708-y.
25. S. Nagpal and V. Ravichandran, Fully starlike and convex harmonic mappings of order α, https://arxiv.org/abs/1207.3946v1.
26. J. K. Prajapat and M. Mathi, Injectivity of Harmonic Mappings with Fixed Analytic Part, Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci. (2022), https://doi.org/10.1007/s40010-022-00781-0
27. A. Raj and S. Nagpal, Radius of Convexity for Analytic part of Sense-Preserving Harmonic Nappings, Bull. Malays. Math. Sci. Soc., (2022). https://doi.org/10.1007/s40840-022-01322-z
28. T. Sheil-Small, Constants for planar harmonic mappings, J. Lond. Math. Soc., 42, 237-248 (1990)
29. Sokół, J.: Radius problems in the class $\mathcal{SL}^*$. Appl. Math. Comput. 214(2), 569–573(2009)

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9836005

Real-Time Object Size Dimensioning in Raspberry Pi

Sahil Khadane
Department of Information Technology
Delhi Technological University
Delhi, India
sahilkhadane_2k18it104@dtu.ac.in

Sahil Saini
Department of Information Technology
Delhi Technological University
Delhi, India
sahilsaini_2k18it105@dtu.ac.in

Samarth Mittal
Department of Information Technology
Delhi Technological University
Delhi, India
samarthmittal_2k18it106@dtu.ac.in

Prof. Kapil Sharma
Head of Department of Information Technology
Delhi Technological University
Delhi, India
kapil@ieee.org

***Abstract*—Real-time object size dimensioning and detection is playing a crucial part in industry today and in the coming days of technological advancement. Object Size Dimensioning is a framework for estimating the dimensions i.e., length and breadth of real-time objects. To demonstrate the efficiency of the suggested system, the demonstration has been implemented and shown in Raspberry Pi. The proposed approach includes an item or object that is identified using improved techniques for detecting edges. For calculating dimension, a reference item (a binary coded marker) has been taken based on characteristics that items used as reference can be identified easily and have fixed measurements. This model has been proposed using the Open CV python module in Raspberry Pi, for determining the dimensioning of an object using a fixed reference.**

Keywords— Raspberry Pi, Open CV, Edge Detection, Object Size Dimensioning, Binary Coded Marker.

I. INTRODUCTION

Recognition system or machine vision is the technique of detecting items in videos and images. These types of IoT devices help in identifying and detecting items of the real world through the computer vision approach. "Real-Time Object Size Dimensioning" or object recognition is a framework for identifying the measurements of real-world objects or use computer vision to recognise items in digital image [1]. The framework works well with an emulator and can be enhanced further to build a functional product using raspberry pi system. Detection and recognition is often utilized in the industry's quality management stages. Developing new concepts linked to machine interface, artificial intelligence, navigation systems in aircraft [2], which includes autonomous vehicles and IoT devices and smart objects which includes sensor activated doors, lights [3]. For eg: An autonomous automobile is a self driving vehicle capable of detecting, sensing, reacting and adapting with respect to its surroundings in order to operate without the support or involvement of a user. Object recognition system in automobiles is seen amongst the most critical problems since it assists the vehicle in tracking objects and deciding its future trajectory [4] uses CNN models [5]. Real-time object for recognition and dimensioning of things is a matter of great interest in a wide range of industries. Traditional and existing methods does not provide much accuracy. It is also a key aspect in machine learning challenges, pattern and surveillance systems for providing security in area such as banks, defence/military or personal security [6].

This project demonstrates a method for calculating measures from images in real time and live video captured with the camera of pi to show overall effectiveness of our proposed framework.

II. LITERATURE REVIEW

Over the years, various techniques for detection of edges were formulated for accurate detection of object contours for the purpose of object detection but each techniques has its own working condition and limitation.

In 1965, Lawrence Roberts developed the technique for the detection of edges known as Robert's edge detection which helps in computing a basic, two dimensional spatial gradient measuring system on a photo. This approach suggested spatial frequency areas, which usually equate with boundaries. The most basic application is via a grayscale image as input into the operator and a grayscale picture as output because of its small kernel it's disadvantage includes weak response and is sensitive to noise [7]. Sobel suggested an enhanced development of Roberts detection in 1970. It measures the two dimensional spatial gradient of a picture, focusing high spatial frequency zones, which generally relate with boundaries. To extract the image's gradient magnitude from within the source, Sobel operator's convolution mask had been implemented. Though it provides good output it is not suitable for complex number of images [8].

Prewitt (1970) provided a useful approach for determining an edge's horizontal and vertical components [9]. Unlike other gradient edge detection methods, which

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:22 UTC from IEEE Xplore. Restrictions apply.

need a lengthy analysis, prewitt detection of edge through compass determines path simply within kernel with greatest throughput response. Though it has a limit of eight possible directions where all eight-kernel masks are produced and then out of the biggest module one complexity mask is chosen [7]. Prewitt detection is somewhat easier to accomplish technically than Sobel detection, but it yields considerably poorer outcomes than any other techniques.

Marr-Hildreth (1980) detection methodology detected contours within pictures where there rapid changes within pixel intensity [8]. To identify edges efficiently, for breaking down the picture whenever the intensity fluctuate, Gaussian filtration and Laplacian are used. Marr introduced LoG which is termed as (Laplacian of Gaussian) for edge detection. A second order derivative [9] which improves picture smoothing and identifies Laplacian, producing a image of double edges, and determines its two thresholds to identify edges, and is defined as f(x,y). In particular, mostly useful for determining how much a pixel is really on either darker and otherwise on brighter side among an edge [7]. It is highly unreliable at curves, contours, and areas where the grey level intensity functions vary.

Through the terms of technological advancement, canny edge detection is still the popular and widely known edge detection technology. Invented by John Canny (1986) at MIT [7], exceeds many algorithms or edge techniques that have been developed recently [10]. Canny is a significantly important method for finding the edge of the object by smoothing and removing its background noise components without damaging the image's edge features and also applying gaussian parameters for determining edge and critical value for threshold [11]. An improved canny edge technique [12] using mathematical morphology [13] was also introduced to make certain variations and introduced new parameters from the existing one because the classic approach is ineffective in dealing with color images [14].

We aim to improve canny edge detection with the help of morphological operations to reduce noise and produce efficient results and implementing the same in Raspberry Pi to work as an efficient working model.

III. WORKING METHODOLOGY

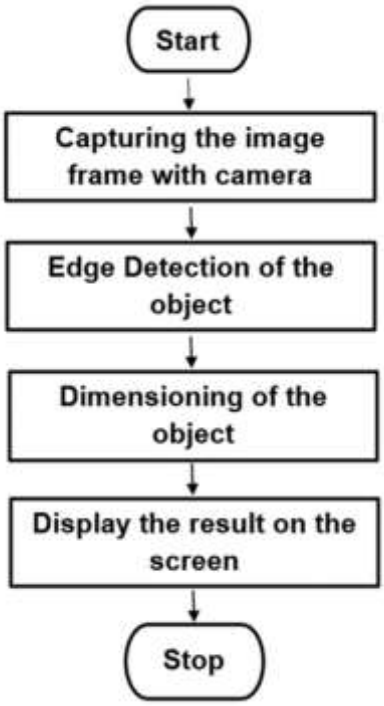


Figure 1. Flow Diagram

Figure 1 depicts all preferred stages together in simplified flow chart. Basic idea was to make a model that determines the height and width of the real time object efficiently with the help of open cv modules and object detection algorithms. The proposed approach includes following steps:

Step 1: items or objects are identified using techniques for detecting edges.
Step 2: morphological operations, dilation and erosion were used to reduce noise and closing spaces between corners and edges.
Step 3: scanning contours and organizing them.
Step 4: computing pixels per metric ratio and converting it into cm.
Step 5: calculating dimension, using a reference item and metric ratio.
Step 6: the Euclidean distance between sets of centre points was then determined and displayed.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:22 UTC from IEEE Xplore. Restrictions apply.

A. Canny Detector

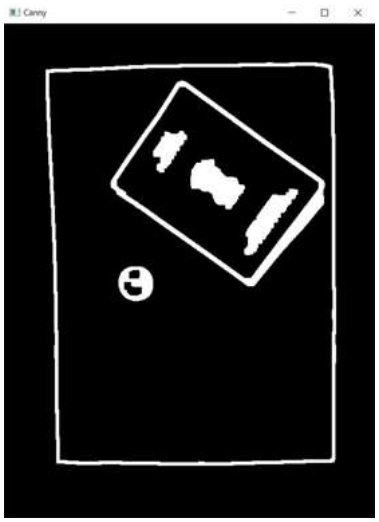


Figure 2. Testing of canny detector

Figure 2 demonstrates objects in the canny frame. The edges detected in Sobel approach are excessively chaotic, and the data has virtually lost its crucial structure. While the other edge techniques either takes too much time or are complex in performing. Thus canny outperforms all the previously suggested edge detection algorithms. Following is the algorithmic approach for canny edge detection.

- Before image's edge is identified, Gauss smoothing filter is used to remove noise. Thus de-noise the image.
- Determine the gradient amplitude and direction, with the four angles of 0, 45, 90, and 135 degrees.
- To position the edges and remove non-edges non-maximum suppression was performed [12].
- We set a hysteresis threshold [10] and to choose the edge, the hysteresis threshold requires two thresholds that keep or exclude pixels. If the edges pixels are stronger than the large threshold the final edges will be classified as strong edges. Whereas if pixel point's amplitude was below threshold, the pixel was eliminated.

B. Dilation and Erosion and use of binary code marker

Structural approaches such as erosion and dilation are examples of morphometric image analysis techniques [15]. Structural image processing using Open CV is a technique for changing a picture's geometrical configuration. The shape, size, and structure of an object are discovered through morphism after Gaussian application [11]. The following are some characteristics of how they are used –

- Eliminating noise.
- Determine whether there are any intensity gaps or holes in the image.
- Individual elements are isolated and dissimilar elements in a picture are joined.
- A binary coded marker of fixed measurement is used for reference to detect the images in the background.

C. Setting up proposed system

Figure 3. Model - Raspberry Pi B Version 1.2

Model B of Raspberry Pi version 1.2 and Pi Camera V2 [16] has been used for implementing the suggested framework to make it a working system shown in figure 3.

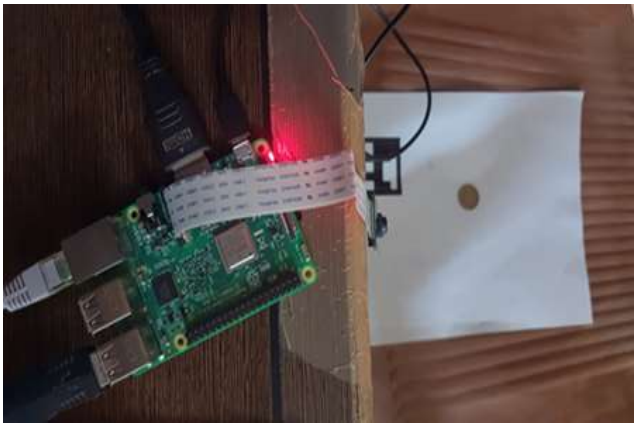


Figure 4. Setting up Raspberry Pi and Camera

After all the connections has been made as seen in figure 4, raspberry pi OS [17] installation is required and then boot up for your first use. The system implementation was done using python language so installation of open cv, its libraries and dependencies are required [18].

D. Detection and Dimensioning

The system executes two main operation first detection of the object and second is dimensioning of the object.

The images of object were captured with pi camera and were transformed it into grayscale. Then Gaussian filter was applied to the converted grayscale image to remove unwanted noise. All the objects placed in the frame were thus captured and after filtering, edges in the picture were detected using canny method. Multiple images can be captured through this edge detection algorithm and can be processed at the same time. Using dilation and erosion also termed as morphological operations were used while

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:22 UTC from IEEE Xplore. Restrictions apply.

filling the gaps present in the edge frame obtained from canny to achieve smoothened frame of canny image. An outline is formed for the objects that were detected using contours command. From left through right, we organize the contours in sequential order. Setting the attribute values based on the reference item. This reference item was uniquely identified in the frame compared to other objects. We have used a binary coded marker and the dimensions of marker was fixed and known to us. To ensure that we have correctly identified the marker, we construct a polygon around it.

After that, we scan all of the contours, to highlight boundaries we used green rectangle layout around items. After obtaining all contours, we calculate pixels, for that we need height and width of reference that is known to us to determine the pixel per metric ratio for which reference item is used. The perimeter was calculated from the detected points and ratio is obtain by division that is used further used for conversion on objects. The height parameters has been set to the height (h) length in pixels, and the width (w) parameter has been set to the width length in pixels. We generate measurement in cm by performing the ratio computation to these two parameters and display the result on screen

E. Experiment Results

The setup of raspberry pi connected with camera was placed over the following first 2 items, a bottle cap and a highlighter. The images were successfully captured by the camera and after the complete execution of the application, dimensions of the items placed were assessed and displayed on the output screen as seen in Figure 5.

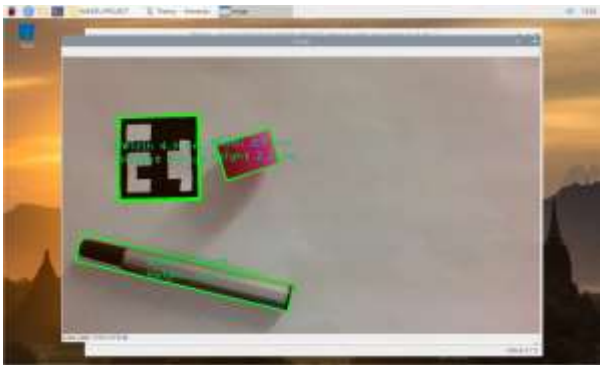


Figure 5. Experiment Screen 1

In the next experiment another 2 items, a coin and a passport were placed below the camera and their dimensions were assessed and thus was displayed on the output screen as seen in Figure 6.

Figure 6. Experiment Screen 2

After the detection and dimensioning of all objects were done through the proposed system, the actual measurements of the items were taken into account to compare with the experimented outcomes to further demonstrate the accuracy estimation of the system.

F. Estimation Table

The system's outcomes were provided and concluded by measuring a significant number of images showing items of different size, and at different distances from the camera. Table 1 demonstrates the suggested object dimensioning system's accuracy estimation for the entities mentioned in the table.

The table's abbreviations are as follows: AH: Actual Height, EH: Estimated Height, AW: Actual Width, EW: Estimated Width.

Name of Object	Actual Dimensions		Estimated Dimensions		Accuracy
	AH (cm)	AW (cm)	EH (cm)	EW (cm)	(%)
Bottle cap	3.3	1.8	3.5	2.2	94.96
Highlighter	1.5	12.9	1.7	13.4	96.12
Coin	2.3	2.3	2.6	2.4	95.6
Passport	12.5	8.8	14.0	9.5	92.04

Table 1. Accuracy Estimation of Objects

The overall observed accuracy of system is approx 94 percent considering the fact that it is not perfect outcome due to various factors such as distance from camera, lens distortion, accuracy prediction of algorithm. Thus, we saw some variations when we compared them with the original dimensions of the items.

IV. CONCLUSION

Through this paper our attempt was to present an object dimensioning model which was achieved with the help of canny edge detection and improved with the use of morphological operators and a binary coded reference marker. The model was successfully tested and executed using raspberry pi and was able to quantify and estimate the dimensions (width and height) of various items in real time video captured. The use of open cv module functions

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:22 UTC from IEEE Xplore. Restrictions apply.

helped in identifying and estimating processes. Our system was able to provide an accuracy of 94 percent. This suggested system is quick, and several images can be captured and analyzed in one frame. The Raspberry Pi has been able to execute the programs because of its robust functionalities thus making it an ideal model making hardware.

Due to the advances and new approaches in computer vision technologies image segmentation, object detection and IOT devices are rising rapidly in various industries for easier and faster detection and recognition. There is much more promising potential and areas for such models to be utilized in the future.

REFERENCES

[1] M. Vashisht and B. Kumar, "A Survey Paper on Object Detection Methods in Image Processing," in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*, Mar. 2020, pp. 1–4. doi: 10.1109/ICCSEA49143.2020.9132871.

[2] S. T. V., M. S, S. R, and B. R, "A Study on Various Mesh Generation Techniques used for Engineering Applications," *Journal of Innovative Image Processing*, vol. 3, no. 2, pp. 75–84, Jun. 2021, doi: 10.36548/jiip.2021.2.001.

[3] I. Aydin and N. A. Othman, "A new IoT combined face detection of people by using computer vision for security application," in *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)*, Sep. 2017, pp. 1–6. doi: 10.1109/IDAP.2017.8090171.

[4] A. Teichman and S. Thrun, "Practical object recognition in autonomous driving and beyond," in *Advanced Robotics and its Social Impacts*, Oct. 2011, pp. 35–38. doi: 10.1109/ARSO.2011.6301978.

[5] Dr. T. S. Kumar, "Video based Traffic Forecasting using Convolution Neural Network Model and Transfer Learning Techniques," *Journal of Innovative Image Processing*, vol. 2, no. 3, pp. 128–134, Jun. 2020, doi: 10.36548/jiip.2020.3.002.

[6] C. Kaundanya, O. Pathak, A. Nalawade, and S. Parode, "Smart Surveillance System using Raspberry Pi and Face Recognition," *IJARCCE*, vol. 6, no. 4, pp. 621–624, Apr. 2017, doi: 10.17148/IJARCCE.2017.64117.

[7] R. Muthukrishnan and M. Radha, "Edge Detection Techniques For Image Segmentation," *International Journal of Computer Science and Information Technology*, vol. 3, no. 6, pp. 259–267, Dec. 2011, doi: 10.5121/ijcsit.2011.3620.

[8] S. Lakshmi and Dr. V. Sankaranarayanan, "A study of Edge Detection Techniques for Segmentation Computing Approaches," *International Journal of Computer Applications*, vol. CASCT, no. 1, pp. 35–41, Aug. 2010, doi: 10.5120/993-25.

[9] Y. Ramadevi, T. Sridevi, B. Poornima, and B. Kalyani, "Segmentation And Object Recognition Using Edge Detection Techniques," *International Journal of Computer Science and Information Technology*, vol. 2, no. 6, pp. 153–161, Dec. 2010, doi: 10.5121/ijcsit.2010.2614.

[10] Z. Xu, X. Baojie, and W. Guoxin, "Canny edge detection based on Open CV," in *2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, Oct. 2017, pp. 53–56. doi: 10.1109/ICEMI.2017.8265710.

[11] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986, doi: 10.1109/TPAMI.1986.4767851.

[12] W. Rong, Z. Li, W. Zhang, and L. Sun, "An improved Canny edge detection algorithm," in *2014 IEEE International Conference on Mechatronics and Automation*, Aug. 2014, pp. 577–582. doi: 10.1109/ICMA.2014.6885761.

[13] Y. Zheng, J. Rao, and L. Wu, "Edge detection methods in digital image processing," in *2010 5th International Conference on Computer Science & Education*, Aug. 2010, pp. 471–473. doi: 10.1109/ICCSE.2010.5593576.

[14] G. Xin, C. Ke, and H. Xiaoguang, "An improved Canny edge detection algorithm for color image," in *IEEE 10th International Conference on Industrial Informatics*, Jul. 2012, pp. 113–117. doi: 10.1109/INDIN.2012.6301061.

[15] J. Y. Gil and R. Kimmel, "Efficient dilation, erosion, opening, and closing algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1606–1617, Dec. 2002, doi: 10.1109/TPAMI.2002.1114852.

[16] Huu-Quoc Nguyen, Ton Thi Kim Loan, Bui Dinh Mao, and Eui-Nam Huh, "Low cost real-time system monitoring using Raspberry Pi," in *2015 Seventh International Conference on Ubiquitous and Future Networks*, Jul. 2015, pp. 857–859. doi: 10.1109/ICUFN.2015.7182665.

[17] "Raspberry Pi OS." https://www.raspberrypi.com/software/ (accessed Mar. 06, 2022).

[18] "OpenCV." https://opencv.org/ (accessed Mar. 06, 2022).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:08:22 UTC from IEEE Xplore. Restrictions apply.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/362814057

Regional Languages Hate Speech Detection System Using IndicBERT and Transfer Learning

Preprint · June 2022

DOI: 10.13140/RG.2.2.29337.70249

CITATIONS
0

READS
12

1 author:

Ayush Tanwar
Delhi Technological University
3 PUBLICATIONS **0** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Folksonomy Based Fuzzy Filtering Recommender System View project

All content following this page was uploaded by Ayush Tanwar on 20 August 2022.

The user has requested enhancement of the downloaded file.

Regional Languages Hate Speech Detection System Using IndicBERT and Transfer Learning

Ayush Tanwar *[1]
[*1] Department of Computer Science & Engineering, Delhi Technological University, India
Keywords: Transformer, Hate Speech Detection, Deep Learning, Natural Language Processing

ABSTRACT
Social media granted people a platform to express their opinions freely. But the anonymity feature of social media allows several individuals to disregard social etiquette to target and hurt specific individuals and groups. In the past few years, Indians have witnessed social unrest, which spread through social media and has injured and taken countless lives. The need of the hour is an advanced deep learning-based hate speech detection system that is not only accurate and efficient but can also work on many regional languages spoken throughout the Indian subcontinent. To manage this need, we introduce an IndicBERT based hate speech detection system that uses Transfer learning and transformers to detect hate speech in online media. To assess our presented research, we use datasets that include hateful, sexist, and offensive tweets. The experimental evaluation reveals that our solution achieves notable improvement over other state-of-the-art models regarding precision and recall. Moreover, our model is designed to collect biases to potentially improve its accuracy over time.

INTRODUCTION
The use of social networking platforms such as YouTube, Twitter, Facebook, etc., to disseminate and share information is growing rapidly. Even though the exchanges among people on these platforms can lead to many positive exchanges, but the propagation of abusive language and the organization of hate-based movements is ever-growing [1]. People take advantage of the ease of accessibility and anonymity on these platforms. Violence due to hate speech propagated online has risen globally. For instance, Asian people faced a lot of hate during the COVID 19 pandemic. Therefore, social media outlets encountering this problem should employ more advanced tools to catch hate speech instances like these, as current mechanisms are not able to do so.

India is a land of hundreds of languages. Traditional hate speech detection systems do not work well with regional Indian languages. As a result, hate speech posts are unnoticed by current systems and succeed at servicing their purpose to spread hate.

In our research, the scope of hate speech is abusive content on social media, as it can be viewed as a general term for various types of offensive user-generated content. Hate speech is commonly defined as a communication attacking an individual or a group based on factors such as sexual orientation, gender, race, religion, nationality, etc. Hate speech identification is not an easy task because misclassification of ordinary exchange as hate speech could negatively affect users' freedom of expression. In contrast, misclassifying offensive exchanges as ordinary would facilitate the tag of online platforms as unsafe territories.

A considerable number of academic researches have been devoted to identify hate speech by utilizing Natural Language Processing (NLP) in a hybrid with Deep Learning (DL) and Machine Learning (ML) methods [2]. Supervised ML-based techniques have utilized various text mining-based traits such as sentiment analysis, surface features, lexical resources, knowledge-based features, and linguistic features; these approaches require a properly-defined feature extraction method. Advancement of Deep learning led to deep learning being used to train classifiers and feature extraction. These hate speech detection methods use deep learning approaches such as Long Short-Term Memory Networks (LSTMs), Convolutional Neural Networks (CNNs)[3], etc., to improve the performance of these models. However, they still suffer from insufficient supervised data and the inability to enhance generalization property.
Hence, a transfer-learning-based approach for hate speech identification employing a mixture of the unsupervised IndicBERT [4] and novel supervised optimization strategies is introduced in this research. It is the foremost hate speech detection model that employs optimization techniques along with a generative pre-trained language model that uses transfer learning for regional hate speech languages and enhances classification accuracy. Rest of the paper is organized in four parts:

1) In Section 2, we talk about other hate speech detection approaches and their advantages and disadvantages
2) In Section 3, We introduce a transfer learning method that uses pre-trained IndicBERT which is trained on large datasets, covering twelve major languages of India. We also introduce novel optimization strategies to analyze the impact of various embedding layers of BERT for offensive speech detection.

3) In section 4, Our experimental results demonstrate that utilizing the pre-trained IndicBERT model and optimizing it on the downstream part by leveraging contextual and syntactical information of all BERT's transformers outperforms other models in terms of metrics like recall, precision, and F1-score. Moreover, analyzing the results reveals the capability of our model to catch biases in annotating or collecting datasets. Future studies could use the bias information collected to improve existing BERT models.
4) In Section 5 We conclude our paper by sharing the findings and discussing future work to extend this research.

RELATED WORK

Hate speech contains a broad range of subjects, and specifying what hate speech exactly means is an onerous task. Anti-Hindu and other such issues are particularly demonstrated to have a rising influence in the Twitter ecosystem. To comprehend how hate speech spreads through Twitter, lexicon-based techniques [5] and belief network techniques [6] are available. Another prominent difficulty of specifying hate speech is choosing what vocabulary correlates to hate speech since the usage of a basic lexicon to detect hate speech can be inaccurate, and the language is constantly changing [7,8]. Our technique, handle this issue by obtaining the context of a word or a group of words using the IndicBERT model. Due to the limitations of lexicon-based approaches, we utilize feature representation of tweets as a whole [9,10]. A broad range of models relating to network embedding have been presented, including node2vec [11], GraphSAGE [12], and skip-gram-based word embedding approach word2vec [13].

Along with such approaches on homogeneous networks, much research has been done on convoluted networks with numerous network layers, edge types, and node types[3]. Established on such complicated connectivity patterns, there are network embedding models that extract features of various nodes in the network, such as multi-net [14], HIN [15], and metapath2vec [16]. In this approach, we utilize such embedding algorithms on a multilayer network of user interactions with interaction scores, and perform node classification task on whether a user is hateful or not.

In [17] writers present techniques for hostile post detection in Hindi. Evaluation of models like CNN, BiLSTM, Multi-CNN, CNN+BiLSTM & mBert, and FastText embedding supplied by both Facebook and IndicNLP. This work displays that BERT-based models work slightly better than basic models. The multi CNN model with IndicNLP FastText word embedding performs best within the basic models.

IndicBERT AND TRANSFER LEARNING BASED HATE SPEECH DETECTOR

In this section, we use the IndicBERT[4] transformer model to detect hate speech. IndicBERT is based on BERT model wich uses multi-layer bidirectional transformer encoder trained on the Book Corpus and the English Wikipedia containing 2,500M and 800M tokens, respectively. IndicBERT has two models named IndicBERTlarge and IndicBERTbase. BERTbase includes an encoder with 12 layers (transformer blocks), 110 million parameters, and 12 self-attention heads, whereas BERTlarge model has 16 attention heads, 24 layers, and 340 million parameters. Calculated embeddings from IndicBERTbase have around 770 hidden dimensions [4].

IndicBERT model is a multi-language ALBERT model which is trained on huge corpora, expanding across 12 main Indian languages: Gujarati, Hindi, Assamese, Kannada, Bengali, English, Malayalam, Punjabi, Tamil, Marathi, Oriya, Telugu. The model has fewer parameters than other models such as XLM-R and mBERT, while it still delivers state-of-the-art performance on various tasks. For our hate speech detection assignment, we are concerned with content on social media, thus as a critical step, we examine the contextual data extracted from IndicBERT' s pre-trained layers and later fine-tune it employing annotated datasets. By tuning we update the weights utilizing a supervised dataset which is new to an pre-trained model. For output and input, IndicBERT uses a sequence of tokens with maximum length 512 and creates an expression of the sequence in a 770-dimensional vector. IndicBERT adds around two segments to every input namely, [SEP] and [CLS]. The [SEP] splits segments and would not be used for classification. [CLS] embedding is the foremost token of the input sequence and possesses the special classification embedding that we use as the first token [CLS] in the last hidden layer for the representation of the entire sequence in the hate speech classification task. To perform the hate speech detection task, we use IndicBERTbase model to classify each tweet as Sexism, Racism, Neither or Offensive and, Hate in our datasets. To accomplish that, we concentrate on fine-tuning the pre-trained IndicBERTbase parameters. By fine-tuning, we imply training a classifier with various layers of 770 dimensions in adddition to the pre-trained IndicBERTbase transformer to reduce task specific parameters.

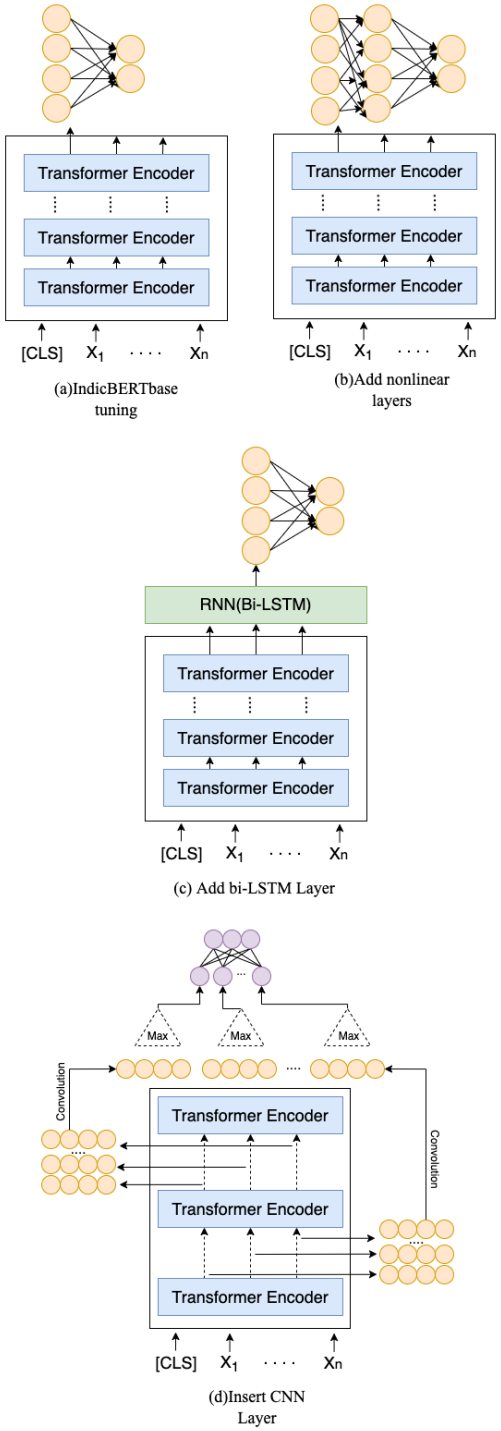


Fig 1: Model Tuning Steps

TUNING IndicBERT

Various layers of a neural network model can catch various levels of semantic and syntactic data. The lower layer of the IndicBERT model may include additional data, but the higher layers include task-specific data [4], and we can fine-tune them with different learning rates. Here, various fine-tuning procedures are executed that use pre-trained IndicBERTbase transformer encoders for the purpose of classification. Additional data about the transformer encoders' architectures are presented in [4]. During fine-tuning phase, the method is started with the pre-trained parameters and then fine-tuned utilizing the supervised datasets. Various fine-tuning approaches for hate speech detection are depicted in Fig. 1, in which X_i is the vector representation of token i for a tweet sample, and are described in further detail in the following:

1. *IndicBERT based fine-tuning:* In the first approach, demonstrated in Fig. 1a, few modifications are done to the IndicBERTbase . Only the [CLS] token's output is provided by IndicBERT is utilized. The [CLS] output, which is equal to the [CLS] token output of the 12th transformer encoder, a vector of size 770, is

provided as input to a completely connected network without a hidden layer. Then activation function softmax is used on the hidden layer to classify.

2. *Add nonlinear layers:* In this step, initially architecture is updated, and the architecture having a better classifier is used. The architechture, instead of utilizing a completely connected network having no hidden layers, a completely connected network with around two hidden layers of size 770 is utilized. Those layers utilize Leaky Relu activation function that has a negative slope of around 0.01. The final layer, as the first design, makes use of softmax activation function depicted in Fig. 1b.
3. *Add Bi-LSTM layer:* previous architectures only used [CLS] as an input for classifier. This architecture contains all outputs of the latest transformer encoder. The outputs are employed as inputs for the bidirectional recurrent neural network depicted in Fig. 1c. After the processing of input, The network sends the last hidden state to a completely connected network that executes classification employing softmax activation function.
4. *Add CNN layer:* The architecture depicted in Fig.1d contains outputs of all transformer encoders used. The output vectors of the individual transformer encoder is concatenated, and an individual matrix is created. The convolutional process is executed using a window of size 3(The discreet length of IndicBERT is 770 in IndicBERTbase model), and a greatest value is generated for individual transformer encoder by using max pooling on the convolutional process output. Connecting these values generates a vector that is provided as input to a completely connected network. By using softmax for the input, the classification task is done.

Table 1. Dataset statistics of the both HOT dataset (a) and MahaHate dataset (b). Splits are done in 0.8: 0.1: 0.1 for training, validation, and testing data, respectively.

	Racist	Sexist	Neither	Total
Training	1303	2,205	4,102	7610
Validation	163	276	513	952
Testing	163	276	513	952
Total	1629	2757	5128	

(a)

	Racist	Sexist	Neither	Total
Training	1881	2,320	20,611	24,812

EXPERIMENTATION & EVALUATION

DATASET DESCRIPTION

The datasets used in our model are the following:

L3Cube- MahaHate[18]

This information is gathered from tweets (26000) from Twitter and then labeled as hate, offensive, profane, and not. It includes Hindi and regional indian language Marathi tweets. We have added English tweets from [20] to test the versatility of our model.

HOT[19]

This dataset includes 7600 tweets categorized into three types which are Non-Offensive, Abusive, and Hate Inducing. Similar to HS dataset, it includes Hindi and English tweets. After data cleaning, the HOT dataset receded to 5100 tweets. We have later added few English tweets from [20] which were not added in MahaHate dataset to test the versatility of our model.

EXPERIMENTATION AND ERROR ANALYSIS

We encounter numbers, usernames, URLs, emoticons & hashtags and replace them with tokens such as <number>, <user>, <url>, <emoticon> & <hashtags>. We also see extended words and transform them to standard and short format; for eg, 'nnnoooo' gets coverted to 'no'. Hashtags that contain tokens without any space between them, gets replaced by their text counterparts; for eg, hashtag “#notaman” gets converted to “not a man”. All punctuation,

unspecified uni-codes and additional delimiting characters were removed. All tweets were converted to lower case.

Method	Dataset	Precision(%)	Recall(%)	F1-Score(%)
MahaHate	MahaHate	91	90	90
HOT	HOT	-	-	80
	MahaHate	-	-	89
IndicBERTbase	HOT	81	81	81
	MahaHate	91	91	91
IndicBERTbase + nonlinear layers	HOT	73	85	76
	MahaHate	76	78	77
IndicBERTbase + LSTM	HOT	87	86	86
	MahaHate	91	92	92
IndicBERTbase + CNN	HOT	89	87	89
	MahaHate	92	92	93

Table 2: Results of testing data using different techniques

We implemented the neural network model using the PyTorch-IndicBERT library comprising the pre-trained IndicBERT model and tokenizer. The implementation is done on the Google Collaboratory tool. Using our learnings from initial experiments, we trained our classifier for three epochs with a batch size of 32. The probability of dropout is fixed to 0.1 for every layer. Adam optimizer having a learning rate of $2e^{-5}$ is employed. For input, each tweet is tokenized using the IndicBERT tokenizer. It contains punctuation splitting, invalid character removal, and changing words to lowercase. Like the BERT model, we divided words to subwords employing WordPiece tokenization. As tweets are generally short, we set the largest sequence size to 64, and if the sequence is shorter, the sequence will be padded with zeros or shortened to maximum length.

We split the dataset using an 80:10:10 split, where 80% would be training data used to revise the parameters in the fine-tuning phase, 10% would be used as validation data for measuring the performance of our model while training, and the rest of the 10% would be used as test data to calculate performance after the training stage. To avoid overfitting, we employ a stratified sampling technique to pick 0.1, 0.1, and 0.8 parts of tweets from every class for validation, testing, and training. The distribution of various classes as validation, testing, and training datasets are demonstrated in Table 1.

As seen in Table 1(a) and (b), The datasets contain class imbalance. As offensive language and hate speech are real-world phenomena, oversampling or under-sampling techniques are not required to adjust the classes' distribution to keep the datasets realistic. We analyze the effect of various tuning techniques on the efficiency of our model. The Table 2 summarizes the received results for tuning techniques along with the original baselines. We take MahaHate [18], and HOT[19] as baselines and compare the findings with various tuning strategies using our pre-trained IndicBERTbase model. The experimentation outcomes are reported on the test dataset using three metrics: recall, precision, and F1-score. We take F1-score as the most rugged metric against class imbalance, which gives us perspicuity into the performance of the proposed techniques. Table 2 shows F1-scores of all IndicBERT-based tuning techniques and shows that the IndicBERT + nonlinear classifier, along with IndicBERT, is more efficient than the baselines. Utilizing the pre-trained IndicBERT model as the initial embeddings and then tuning the model with a completely connected linear classifier (IndicBERTbase) surpasses other baselines delivering an F1-score of 82% and 92% for datasets HOT and MahaHate. Adding a CNN to the pre-trained IndicBERT model for tuning on the downstream stage delivers the most suitable results of F1-score of 89% and 93% for datasets HOT and MahaHate which surpasses the baselines. Consequently, it makes sense that uniting every pre-trained IndicBERT layer with a Convoluted Neural Network produces more promising results. Our technique utilizes all the information contained in various layers of pre-trained IndicBERT during the tuning stage. This information includes syntactical and contextual features ranging from lower to higher layers of BERT.

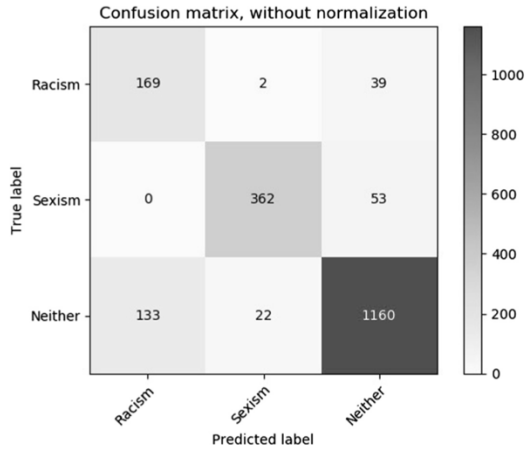


Fig 2: HOT dataset confusion matrix

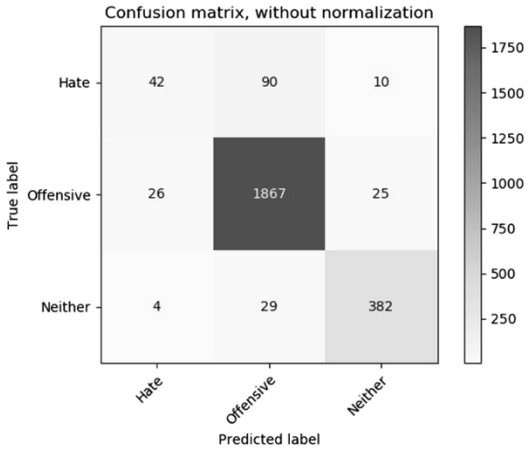


Fig 3: MahaHate dataset confusion matrix

Table 3: Misclassified samples from HOT

ID	Tweet	Annotated	Predicted
1	@user Good tweet. But they actually start selling their daughters at 9	Racism	Sexism
2	RT @user: Are we going to continue seeing the oppression of women or are we going to make a stand? #BanTheBurka	Racism	Sexism
3	RT @user: @user my comment was sexist, but I'm not personally, always a sexist	Sexism	Neither
4	RT @user: @user Ah, you're a #feminist? Seeing #sexism everywhere then, do check my tweets before you call me	Sexism	Neither
5	@user By hating the ideology that enables it, that is what I'm doing	Racism	Neither

Table 4: Misclassified samples from MahaHate

ID	Tweet	Annotated	Predicted
6	@user: If you claim Macklemore is your favorite rapper I'm also assuming you watch the WNBA on your free time faggot	Hate	Offensive
7	@user: Some black guy at my school asked if there were colored printers in the library. "It's 2014 man you can use any printer you want" I said	Hate	Neither
8	RT @user: @user typical coon activity	Hate	Neither
9	@user: @user @user White people need those weapons to defend themselves from the subhuman trash your sort unleashes on us	Neither	Hate
10	RT @user: Finally! Warner Bros. making superhero films starring a woman, person of color and actor who identifies as ""queer"";	Neither	Offensive

Even though we got promising outcomes for recall, the precision results of the technique reveal the amount of incorrect detection we deliver. To better understand this, in this part, we conduct an in-depth study on the various error in the model. We analyze the test datasets and the confusion matrices we get from the IndicBERTbase + CNN model as the most suitable tuning technique, represented in Figs.2 and 3. The Fig.2 for HOT-dataset shows that the technique can distinguish racism content from sexism content adequately. Only a few samples of racist class were misclassified as sexist, and no sexist instances are misclassified as racist. A considerable number of errors came from misclassifying offensive categories (sexism and racism) as inoffensive (neither) and vice versa. 0.8% and 17.5% of the racist instances were misclassified as sexist and neither, respectively, whereas 0% and 12.8% for sexist samples. Nearly 13% of neither samples were misclassified as sexist or racist. Figure 3 shows in MahaHate-dataset that the bulk of fallacies are associated with hate class, in which our technique misclassified hatred tweets as offensive in 64% of the instances. Nonetheless, 3.6% and 8.9% of neither and offensive samples were misclassified.

To better understand misclassified items by our technique, we do a manual assessment on a subset of data and document it in Tables 3 and 4. Words like "women", "daughters", and "burka" are used in tweets with the IDs 1 and 2 in Table 3, our IndicBERT based model is confounded with the contextual meaning between these words

in such samples and misclassified them as sexist, as they are mainly related to femininity. In many cases, including implied abuse such as IDs 5 and 7, our method cannot detect the offensive content and hence misclassifies.

By studying more examples and concerning recent studies, it is evident that many errors exist due to biases from data collection and annotation rules and not because of classifier itself. As HOT [19] used a small ad-hoc set of keywords and MahaHate [18] used a sizeable crowd-sourced lexicon of keywords for sampling tweets for training. The datasets included biases in the gathered data.

Misclassifications do not ensure the low accuracy of our classifier as annotators tend to annotate numerous samples possessing impolite words as offensive without any presumption regarding the social context of users, such as the speaker's name or the dialect. In contrast, they were just neither or offensive tweets. Tweets with IDs 6, 8, and 10 are few samples including offensive words and insults that are not offensive in every case, and tweeters of them use this sort of language in their everyday communications. With these components of evidence, by assessing the content of tweets, we can see in tweets IDs 3, 4, and 9 that our InidcBERT-based classifier can distinguish tweets in which implied hatred content exists and neither exists. One reason for this behavior may be the pre-trained general knowledge which exists in our technique. As the pre-trained IndicBERT model is trained on 12 regional languages corpora, it has learned general knowledge from textual data without purposely offensive or hateful language. Thus, despite the bias, our technique can accurately distinguish offensive and hate samples by utilizing knowledge-aware speech understanding. It can be the primary cause for heightened misclassifications of offensive samples as hate.

RESULTS AND FUTURE WORK

Commixing hate-inciting tweets with benign language lead online hate speech detection mechanisms to detect user-generated content wrongly. Not handling this issue may lead to extremely negative consequences for social media services and their users, such as harm to reputation or user boycott. Hence, we present a transfer-learning-based technique using the pre-trained speech model IndicBERT to improve the performance of a hate speech detection model and make it adaptable to various datasets. We accomplish that by introducing novel tuning methods to analyze the impact of multiple layers of IndicBERT for hate speech detection. The experimentation results show that by employing a CNN-based tuning technique, our technique surpasses previous techniques by benefiting from the contextual and syntactical information embedded in various transformer encoder layers of the IndicBERT model. Moreover, analyzing the results indicates the capability of our technique to catch some biases in the process of annotating or collecting datasets. This can be valuable as it can be used in the pre-trained IndicBERT model to ameliorate bias in offensive speech datasets in future research. In the future, we plan to employ this technique for code-switched low-resource languages.

REFERENCES

1. Yin, Wenjie, and Arkaitz Zubiaga. "Towards generalisable hate speech detection: a review on obstacles and solutions." *PeerJ Computer Science* 7 (2021): e598.
2. Mullah, Nanlir Sallau, and Wan Mohd Nazmee Wan Zainon. "Advances in machine learning algorithms for hate speech detection in social media: a review." *IEEE Access* (2021).
3. Van Huynh, Tin, et al. "Hate speech detection on vietnamese social media text using the bi-gru-lstm-cnn model." *arXiv preprint arXiv:1911.03644* (2019).
4. Kakwani, Divyanshu, et al. "IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages." *Findings of the Association for Computational Linguistics: EMNLP 2020*. 2020.
5. Mathew, B., Dutt, R., Goyal, P., Mukherjee, A.: Spread of hate speech in online social media. In: ACM Web Science, pp. 173–182 (2019)
6. Ribeiro, M.H., Calais, P.H., Santos, Y.A., Almeida, V.A., Meira Jr, W.: Characterizing and detecting hateful users on twitter. In: AAAI ICWSM (2018)
7. Davidson, T., Warmsley, D., Macy, M., Weber, I.: Automated hate speech detection and the problem of offensive language. In: AAAI ICWSM (2017)
8. Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., Chang, Y.: Abusive language detection in online user content. In: ACM WWW, pp. 145–153 (2016)
9. Djuric, N., Zhou, J., Morris, R., Grbovic, M., Radosavljevic, V., Bhamidipati, N.: Hate speech detection with comment embeddings. In: ACM WWW, pp. 29–30 (2015)
10. Le, Q., Mikolov, T.: Distributed representations of sentences and documents. In: ICML, pp. 1188–1196 (2014)
11. Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks. In: ACM SIGKDD, pp. 855–864 (2016)
12. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: NIPS, pp. 1024–1034 (2017)

13. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: NIPS, pp. 3111–3119 (2013)
14. Bagavathi, A., Krishnan, S.: Multi-Net: a scalable multiplex network embedding framework. In: Aiello, L.M., Cherifi, C., Cherifi, H., Lambiotte, R., Li´o, P., Rocha, L.M. (eds.) COMPLEX NETWORKS 2018. SCI, vol. 813, pp. 119–131. Springer, Cham (2019)
15. Dong, Y., Chawla, N.V., Swami, A.: metapath2vec: scalable representation learning for heterogeneous networks. In: ACM SIGKDD, pp. 135–144 (2017)
16. Shi, C., Hu, B., Zhao, W.X., Philip, S.Y.: Heterogeneous information network embedding for recommendation. IEEE Trans. Knowledge Data Eng. 31(2), 357– 370 (2018)
17. Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., Chang, Y.: Abusive language detection in online user content. In: ACM WWW, pp. 145–153 (2016)
18. Velankar, Abhishek, et al. "L3Cube-MahaHate: A Tweet-based Marathi Hate Speech Detection Dataset and BERT models." *arXiv preprint arXiv:2203.13778* (2022).
19. A. Bohra, D. Vijay, V. Singh, S. S. Akhtar, M. Shrivastava, A dataset of hindi-english code-mixed social media text for hate speech detection, in: Proceedings of the second workshop on computational modeling of people's opinions, personality, and emotions in social media, 2018, pp. 36–41.
20. Davidson, T., Warmsley, D., Macy, M.W., et al.: Automated hate speech detection and the problem of offensive language. CoRR abs/1703.04009 (2017). http:// arxiv.org/abs/1703.04009

View publication stats

Scientific Journal of Silesian University of Technology. Series Transport
Zeszyty Naukowe Politechniki Śląskiej. Seria Transport

Volume 115 **2022**

p-ISSN: 0209-3324

e-ISSN: 2450-1549

DOI: https://DOI:.org/10.20858/sjsutst.2022.115.11

Journal homepage: http://sjsutst.polsl.pl

Article citation information:
Sharma, M.K. Cloud manufacturing: identifications and prioritization of opportunities using AHP. *Scientific Journal of Silesian University of Technology. Series Transport*. 2022, **115**, 161-173. ISSN: 0209-3324. DOI: https://DOI:.org/10.20858/sjsutst.2022.115.11.

Mohneesh Kumar SHARMA[1]

CLOUD MANUFACTURING: IDENTIFICATIONS AND PRIORITIZATION OF OPPORTUNITIES USING AHP

Summary. The origin of a cost-efficient, service-oriented, customer-centric, manufacturing system called cloud manufacturing has evolved due to advancements in cyber systems and the availability of internet facilities worldwide. However, there is a significant number of opportunities before the adoption of cloud manufacturing. Through literature survey, expert opinions from academicians and industrialists, various opportunities, namely, pay-as-use, scalability, cost efficiency, flexibility, autonomy, low-risk backup and recovery, low startup cost and location independence associated with the espousal of cloud manufacturing are identified. Further, the Analytic Hierarchy Process (AHP) model is applied to find the weights and prioritize these opportunities, thereby finding the significant key opportunities. Moreover, the consistency ratio is calculated for the accuracy and consistency of the results. As the obtained value of consistency ratio is less than .1, it shows that the result obtained is consistent and accurate. The managerial implication of these outcomes is that the results would indirectly help entrepreneurs in the adoption of cloud manufacturing.

Keywords: cloud manufacturing, opportunities, pay-per-use, scalability, analytic hierarchy process

[1] Department of Mechanical Engineering, Delhi Technological University, New Delhi 110042, India. Email: simplymohneesh@gmail.com. ORCID: https://orcid.org/0000-0003-3347-9700

1. INTRODUCTION

Recent advancements in technologies and customized requirements of customers and manufacturing industries have gone through significant changes leading to a highly competitive environment globally. Cost, customization, flexibility, reliability, and quality are the key factors that enterprises must improve to sustain in the global market. Following the path of networking some of the manufacturing models were introduced, namely, Agile Manufacturing (AM) [1, 2], Network Manufacturing (NM) [3, 4], Manufacturing Grid (MG) [5, 6], Virtual Manufacturing (VM) [7, 8], Additive Manufacturing [9, 10], and Smart Manufacturing (SM) [11, 12], which have changed the manufacturing process significantly. The cloud computing model helps in enabling a user secure and convenient on-demand network access to a shared pool of computing resources like storage, network, applications services, etc. [13]. Moreover, the virtualization and service-oriented characteristics of cloud computing make manufacturing suitable for customization. Gibson et al. (2012) explained the benefits of service models of cloud computing [14]. When services like SaaS, PaaS, and IaaS of cloud computing are applied to manufacturing, then the new manufacturing model obtained is called Cloud Manufacturing [15].

Cloud manufacturing is preferred over other manufacturing systems due to its cost efficiency, pay as per requirement and scalability. Ghomi et al. (2019) provided an overview of cloud manufacturing challenges, recent advancements, issues related to research, and its future trend [16].

Since this manufacturing model is growing at a significant pace, its merits should not be left unknown. Knowledge regarding positive aspects of the manufacturing model helps in its future development and eases its adoption. Narwane et al. (2019) presented a review on issues related to manufacturing and its adoption, they also highlighted the application of the manufacturing model for various industries and sectors [17]. Abubakar et al. (2014) studied the issues related to the adoption of cloud computing for small and medium-sized enterprises in developing countries of the Saharan Africa [19]. Zhang et al. (2020) developed a service platform to increase the competitiveness among small and medium scale industries regarding cloud manufacturing [21]. Abd et al. (2018) compared the current adoption of the manufacturing model with the ideal manufacturing model [22]. Brief studies have been done in past papers on the merits of cloud manufacturing; however, no study deals with its prioritization. Hence, this endeavour is made to identify the opportunities of cloud manufacturing and their prioritization.

Cloud manufacturing is based on shared manufacturing infrastructure, services, and resources through a cloud platform. It uses algorithms to make intelligent decisions and provides the most optimized way for sustainable and robust manufacturing [23]. The recent advancements in technology have made manufacturing more flexible, resourceful, efficient, and customized. Given that cloud manufacturing is still an explored field in developed countries and a new concept to be adopted by developing countries, this study aims to identify the benefits of cloud manufacturing for improving the existing model of cloud manufacturing in developed countries and the initial adoption of cloud manufacturing model in developing countries. Further, it is to identify the opportunities of the cloud manufacturing model and prioritize them to provide prior information on the benefits of this manufacturing model to entrepreneurs interested in its adoption, especially in developing countries. This paper mainly discusses the opportunities of cloud manufacturing. Hence, identification of opportunities is done and through AHP key opportunities are obtained.

The brief of this paper is as follows:
1. Identification of cloud manufacturing opportunities.
2. Application of the AHP approach for prioritization of opportunities.
3. Analysis, result, discussion and conclusion.

2. OPPORTUNITIES

In this section, the opportunities of cloud manufacturing is discussed. Significant factors supporting the adoption of cloud manufacturing are pay-per-use, scalability, cost efficiency, flexibility, autonomy, low-risk backup and recovery, low startup cost, and location independence. All the opportunities are discussed in the section below and finally listed in Table 1.

2.1. Pay-Per-Use or Pay-As-You-Go Service

A user can request services as and when required and can pay according to the "pay-per-use" model, where he pays for the time he has availed of the services, resources, or infrastructure. This scheme is managed by a cloud platform [23]. Pay-as-you-go facility can be availed by users without having any direct interaction with the service provider [24]. The pay-as-you-go model ensures the exchange of services between the manufacturing providers and consumers [25]. It also promotes the paying scheme named pay-as-you-go to the customers [26, 27].

2.2. Scalability

Cloud manufacturing facilitates the user to run a production system on market demand [23]. It provides the facility of scaling where the user can scale (up or down) the use of resources according to his needs [24]. Removal, modification, and addition of resources can be done as required [25]. Cloud manufacturing makes it easier to scale up or down their production according to customer demand [28].

2.3. Cost Efficiency

The support of the Internet, IoT (Internet of things), and Big Data to manufacturing lowers the entry cost for smaller firms. Moreover, with the optimized use of resources, this manufacturing system has become the most cost-efficient manufacturing system in recent times [29]. Cloud manufacturing leads to an increase in the usage of manufacturing resources through outsourcing [25]. By adopting cloud manufacturing, the cost of manufacturing can be reduced [30].

2.4. Flexibility

Cloud manufacturing can adapt to unpredicted changes in circumstances [23]. Further, it can also adapt and respond to changing customer demands [25]. Cloud manufacturing can generate new types of classes of manufacturing processes and deliver manufactured products to clients [28].

2.5. Autonomy

Every manufacturing provider and customer is an independent identity and works independently without having a direct link between them. For cost reduction in operating costs, considerable autonomy is needed in the administration [31]. Service agents provide this autonomy [32].

2.6. Low-risk Backup and Recovery

Cloud manufacturing reduces the risk for small-scale enterprises as all the tasks are outsourced to other companies through the cloud platform [27]. Data storage by the provider makes the data safe with a recovery functionality in case of an emergency. Sharing the benefits among manufacturers lowers the risk factor [30].

2.7. Low Startup Cost

Being that there is no infrastructure or machinery, and the user owns the software, working on the pay-as-use model makes it free from upfront investment for start-ups. Cloud Manufacturing lowers startup and operating costs [27]. Infrastructure and administration costs are also reduced resulting in lower upgrading and maintenance costs [28].

2.8. Location Independence

Independence from locational constraints for the user and the provider takes the freedom of this system to another level. Work can be performed anywhere and at any time. Tasks can be done from suitable enterprises located at any place, thus making this model location independent. Since the system is independent of location, the customer is not required to be concerned about the location of the resources he is using [33]. There is a sense of locational independence where the customer has no control over the location of the service provider [34].

Tab. 1.

Opportunities in espousal cloud manufacturing

S.No.	Opportunities	References
O_1	Pay-per-use or Instant service	[23-27]
O_2	Scalability	[23,24,25,28]
O_3	Cost efficiency	[25,29,30]
O_4	Flexibility	[23,25,28]
O_5	Autonomy	[31,32]
O_6	Low-risk backup and recovery	[27,30]
O_7	Low startup cost	[27,28]
O_8	Location independence	[33,34]

3. RESEARCH METHODOLOGY

In this article, the Analytic Hierarchy Process (AHP), one of the popular MCDM techniques, is used to prioritize identified opportunities in the adoption of cloud manufacturing. A consistency ratio to identify whether the results obtained are robust and consistent is obtained for the validation. For the stepwise application of this method, the flowchart is displayed in Figure 1. AHP can be applied to many sectors in the industry for selection making, decision making, and prioritization [35]. Hu et al. (2019, 2020) used it in the selection of manufacturers in cloud manufacturing [36, 37]. Sevinc et al. (2018) used to solve issues SMEs face transitioning to Industry 4.0. Mian et al. (2020) applied SWOT-AHP to quantify and rank the opportunities and challenges for sustainability education in Industry 4.0. While Metin Da˘gdeviren (2008) used the AHP–PROMETHEE integrated approach for decision making in equipment selection [38-40]. Prioritization of challenges to Industry 4.0 for the supply chain is obtained with the help of AHP [41]. The cloud manufacturing concept is quite new so finding experts in this field is difficult; however, a total of 30 experts from the industry and academics (10 industrialists, 10 mechanical engineering academicians, 6 computer science academicians, and 4 industrial engineering academicians) were used for this study. After discussion, a total of 8 opportunities were finalized and detailed questionnaires (Appendix-1) were sent to the experts for filling. After collecting the responses, an average was obtained to get the final average pairwise matrix.

Steps to apply the AHP approach are as follows:

Step 1: Develop a structural hierarchy

Step 2: Develop an average pairwise comparison matrix

Experts were asked to fill the survey questionnaires where they have to do a pairwise comparison of attribute i with attribute j on a scale of 1, 3,5,7,9. Then all questionnaire matrices are collected and the average matrix Aij is obtained.

$$Aij= \begin{bmatrix} a_{11} \ldots\ldots a_{1j} \ldots\ldots a_{1n} \\ a_{i1} \ldots\ldots a_{ij} \ldots\ldots a_{1j} \\ a_{n1} \ldots\ldots a_{nj} \ldots\ldots .a_{nn} \end{bmatrix}$$

Step 3: Develop normalized decision matrix

$cij = aij / \sum_{j=1}^{n} aij$ (1)

where i=1,2,3,4……. n and j=1,2,3,4……. n

Step 4: Develop weighted normalized decision matrix

$w_i = \sum_{j=1}^{n} cij$ where i=1,2,3,4……. N (2)

Step 5: Calculate eigenvector and row matrix

$E = N^{th}$ rootvalue $/ N^{th}$ rootvalue (3)

Row matrix= $\sum_{j=1}^{n} aij * ej1$ (4)

Step 6: Calculate the maximum eigenvalue, λ_{max}.

λ_{max} =Rowmatrix / E (5)

Step 7: Obtain the Consistency Index (CI) and Consistency Ratio (CR).

CI= (λ_{max} - n)/ (n-1) (6)

CR=CI/RI (7)

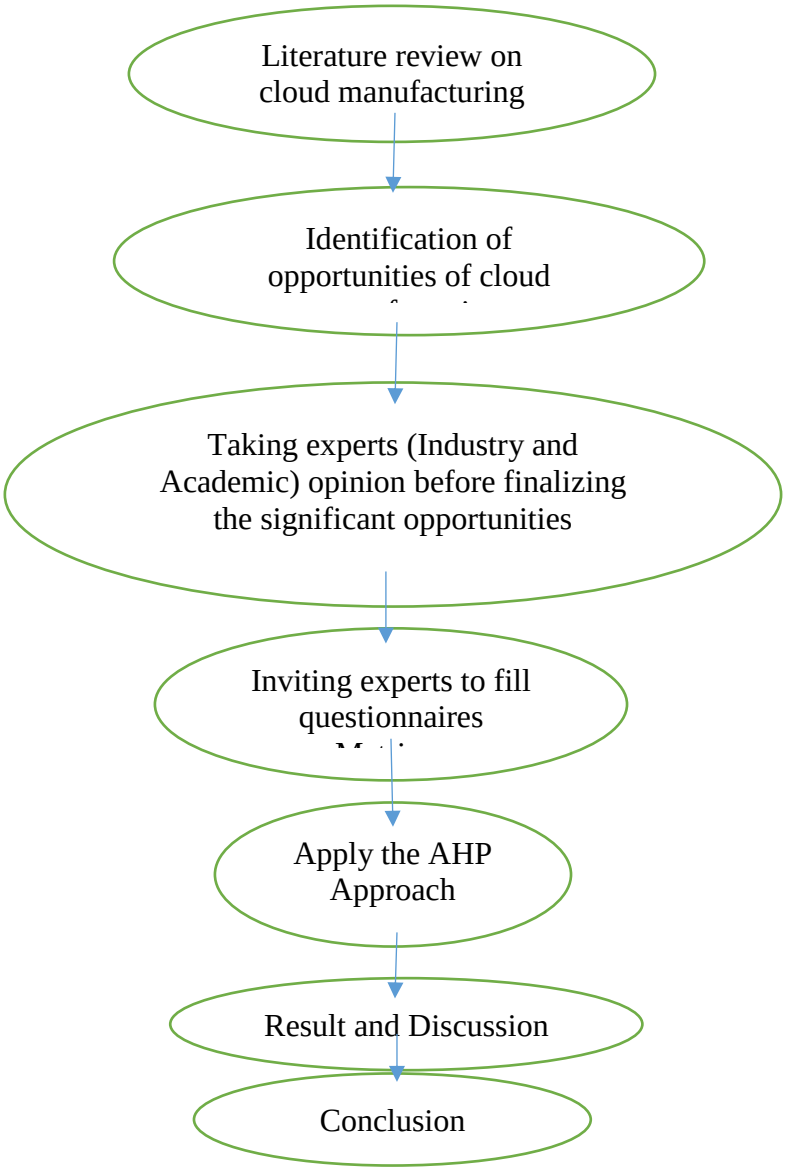


Fig. 1. Research methodology flowchart

Tab. 2.

Random Index

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.85	0.9	1.12	1.24	1.32	1.41	1.45	1.51

Where n and RI denote the order of matrix and Randomly Generated Consistency Index, respectively.

4. RESULTS AND DISCUSSION

This section discusses the prioritization of opportunities and the consistency of results obtained. The consistency ratio is obtained for the validation of the result. After averaging all the individual matrices obtained from the experts, a final average pairwise comparison matrix is obtained followed by the procedure of the AHP approach given in section 3. For all calculation purposes, the matrices are linked in Microsoft Excel to obtain error-free and accurate results.

After the average pairwise comparison matrix is obtained for opportunities as shown in Table 3, normalization is done to get the normalized decision matrix (Table 4). Thereafter, further steps are followed regarding the AHP approach to obtain a weighted normalized decision matrix (Table 5) for weights of opportunities. Finally, ranking is done based on weights obtained (Table 6).

For the final result obtained, Cost-efficient (O_3) has the highest positive value; therefore, it is the most critical opportunity in this category. Pay-per-use (O_1), Scalability (O_2), and Flexibility (0_4) secured the 2nd, 3rd, and 4th places, respectively, in the category as per their weights obtained. Low-risk Backup and Recovery (O_6) and Low Startup Cost (O_7) are ranked 5th and 6th; however, since the differences in weights are very less so they can be assumed at the same level if required. Location Independence (O_8) factor obtained the lowest weight and is ranked 8th among all. All the above rankings are shown in Table 6.

Table 7 is used to calculate the consistency ratio. The obtained value of λmax (Table 7) and random index (Table 2) is finally used to calculate the value of Consistency Ratio (CR). The obtained value of CR is .07, which is less than .10, implying that the result is accurate and consistent.

Tab. 3.

Average pairwise comparison matrix

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8
O_1	1.00	1.67	0.29	3.00	6.33	5.00	5.00	6.33
O_2	0.78	1.00	0.29	3.00	5.00	3.00	3.00	6.33
O_3	3.67	3.67	1.00	3.67	5.67	5.00	5.00	6.33
O_4	0.33	0.33	0.33	1.00	3.00	3.00	3.00	5.67
O_5	0.16	0.20	0.20	0.33	1.00	0.29	0.29	3.00
O_6	0.20	0.33	0.20	0.33	3.67	1.00	1.67	3.00
O_7	0.20	0.33	0.20	0.33	3.67	0.78	1.00	3.00
O_8	0.16	0.15	0.15	0.20	0.33	0.33	0.33	1.00

Tab. 4.

Normalized decision matrix

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8
O_1	0.154	0.217	0.108	0.253	0.221	0.272	0.259	0.183
O_2	0.119	0.130	0.108	0.253	0.174	0.163	0.156	0.183
O_3	0.564	0.478	0.377	0.309	0.198	0.272	0.259	0.183

O_4	0.051	0.043	0.125	0.084	0.105	0.163	0.156	0.163
O_5	0.025	0.026	0.075	0.028	0.035	0.016	0.015	0.087
O_6	0.031	0.043	0.075	0.028	0.128	0.054	0.086	0.087
O_7	0.031	0.043	0.075	0.028	0.128	0.042	0.052	0.087
O_8	0.025	0.019	0.055	0.017	0.012	0.018	0.017	0.029

Tab. 5.

Weighted normalized decision matrix

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	Weighted sum value	Weight
O–1	0.154	0.217	0.108	0.253	0.221	0.272	0.259	0.183	1.667	0.2084
O–2	0.119	0.130	0.108	0.253	0.174	0.163	0.156	0.183	1.287	0.1609
O–3	0.564	0.478	0.377	0.309	0.198	0.272	0.259	0.183	2.640	0.3300
O–4	0.051	0.043	0.125	0.084	0.105	0.163	0.156	0.163	0.889	0.1112
O–5	0.025	0.026	0.075	0.028	0.035	0.016	0.015	0.087	0.306	0.0383
O–6	0.031	0.043	0.075	0.028	0.128	0.054	0.086	0.087	0.532	0.0665
O–7	0.031	0.043	0.075	0.028	0.128	0.042	0.052	0.087	0.486	0.0607
O–8	0.025	0.019	0.055	0.017	0.012	0.018	0.017	0.029	0.192	0.0240

Tab. 6.

Ranking matrix

	Weight	Rank
O_1	0.2084	2
O_2	0.1609	3
O_3	0.3300	1
O_4	0.1112	4
O_5	0.0383	7
O_6	0.0665	5
O_7	0.0607	6
O_8	0.0240	8

Tab. 7.

Calculation of consistency

	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	Weighted value (WV)	Weight (W)	R=WV/W
O_1	1.00	1.67	0.29	3.00	6.33	5.00	5.00	6.33	1.92	0.2084	9.231725
O_2	0.78	1.00	0.29	3.00	5.00	3.00	3.00	6.33	1.46	0.1609	9.096636
O_3	3.67	3.67	1.00	3.67	5.67	5.00	5.00	6.33	3.10	0.3300	9.391174
O_4	0.33	0.33	0.33	1.00	3.00	3.00	3.00	5.67	0.97	0.1112	8.754888
O_5	0.16	0.20	0.20	0.33	1.00	0.29	0.29	3.00	0.31	0.0383	8.092786
O_6	0.20	0.33	0.20	0.33	3.67	1.00	1.67	3.00	0.57	0.0665	8.555815
O_7	0.20	0.33	0.20	0.33	3.67	0.78	1.00	3.00	0.51	0.0607	8.468272
O_8	0.16	0.15	0.15	0.20	0.33	0.33	0.33	1.00	0.21	0.0240	8.595446

$\lambda max=8.77$ RI=1.41

CI= (8.77-8)/7 = .11 CR=.11/1.41= .07

In the values obtained above, λmax is the eigenvalue of the final matrix obtained from Table 7. RI is a random index obtained from Table 2 as the number of factors is 8, thus corresponding to the number 8 value, 1.41, is taken for calculation. Consistency index (CI) is obtained by applying equation 6. Finally, CR is the ratio of consistency index to random index, which signifies how much the observed values and the calculated values are related, and by applying equation 7 and it came out to .07, which means the observed value is very close to the calculated value, so the result obtained is accurate.

As costs remain the most important criteria for almost all the manufacturing systems, therefore, from the result obtained in section 4, Cost-efficient (O_3) emerged as the most critical opportunity in this category. While Pay-per-use (O_1) became the next most important opportunity as paying heavy amounts at one time for small or medium scale industries is critical; therefore, the facility to pay as per requirement attracts them. Scalability (O_2), the facility of scaling where the user can scale (up or down) the use of resources according to his needs secured the 3rd place. Flexibility (0_4), the ability to adapt and respond to the changing customer demands secured the 4th place. Lastly, the Location Independence (O_8) factor is the least important criterion on the list.

5. CONCLUSION

To provide easy decision-making for entrepreneurs for the adoption of cloud manufacturing, opportunities are identified and prioritization of these parameters is obtained. Identification of these opportunities was obtained by experts' opinions and a survey was conducted to get raw data from experts from different fields, namely, industrial, computer, and mechanical and also from professions such as the academics and the industry to obtain the ranking using the MDCM

method. Identified opportunities in the context of cloud manufacturing are, namely pay-as-use, scalability, cost efficiency, flexibility, autonomy, low-risk backup and recovery, low startup cost, and location independence. With the application of AHP on data obtained from the survey, weights are calculated and ranking is obtained. Results show that cost efficiency is the biggest opportunity in the adoption of cloud manufacturing. Furthermore, the calculated consistency ratio for opportunities is .07, which is less than .10, indicating accuracy and consistency of results. Cloud manufacturing process originated in China and spread to countries like Japan, the U.S, Canada, Germany, France Russia, and some countries in Latin America. This study provides prior information regarding the significant factors of the process to entrepreneurs of developing countries, especially for small-scale industries for easy and smooth adoption of the cloud manufacturing process. Conclusively, process cost efficiency, scalability, and pay-as-use qualities would surely attract new industrialists for its setup in developing countries.

References

1. Gunasekaran Angappa. 1999. "Agile manufacturing: a framework for research and development". *International journal of production economics* 62(1-2): 87-105.
2. Sindhwani Rahul, Vasdev Malhotra. 2017. "A framework to enhance agile manufacturing system: a total interpretive structural modelling (TISM) approach". *Benchmarking: An International Journal*.
3. Cheng Yang, Sami Farooq, John Johansen. 2011. "Manufacturing network evolution: a manufacturing plant perspective". *International Journal of Operations & Production Management*.
4. Cheng Tao, Jie Zhang, Chunhua Hu, Bo Wu, Shuzi Yang. 2001. "Intelligent machine tools in a distributed network manufacturing mode environment". *The International Journal of Advanced Manufacturing Technology* 17(3): 221-232.
5. Fan Yushun, Dazhe Zhao, Liqin Zhang, Shuangxi Huang, Bo Liu. 2003. "Manufacturing grid: needs, concept, and architecture". In: *International conference on grid and cooperative computing*: 653-656. Springer, Berlin, Heidelberg.
6. Tao Fei, Ye Fa Hu, Zu De Zhou. 2008. "Study on manufacturing grid & its resource service optimal-selection system". *The International Journal of Advanced Manufacturing Technology* 37(9): 1022-1041.
7. Shukla Chetan, Michelle Vazquez, Frank Chen. 1996. "Virtual manufacturing: an overview". *Computers & Industrial Engineering* 31(1-2): 79-82.
8. Lee Wing Bun, Chi Fai Cheung, J. G. Li. 2001."Applications of virtual manufacturing in materials processing". *Journal of materials processing technology* 113(1-3): 416-423.
9. Guo Nannan, Ming C. Leu. 2013. "Additive manufacturing: technology, applications and research needs". *Frontiers of mechanical engineering* 8(3): 215-243.
10. Wong Kaufui V., Aldo Hernandez. 2012. "A review of additive manufacturing". *International scholarly research notices* 2012.
11. Kusiak Andrew. 2018. "Smart manufacturing". *International Journal of Production Research* 56(1-2): 508-517.
12. Wang Jinjiang, Yulin Ma, Laibin Zhang, Robert X. Gao, Dazhong Wu. 2018. "Deep learning for smart manufacturing: Methods and applications". *Journal of manufacturing systems* 48: 144-156.

13. Mell Peter, Tim Grance. 2011. "The NIST definition of cloud computing". Recommendations of the National Institute of Standards and Technology. NIST National Institute of Standards and Technology U.S. Department of Commerce. *Special Publication 800-145.*
14. Gibson Joel, Robin Rondeau, Darren Eveleigh, Qing Tan. 2012. "Benefits and challenges of three cloud computing service models". In: *2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN)*: 198-205. IEEE.
15. Xu Xun. 2012. "From cloud computing to cloud manufacturing". *Robotics and computer-integrated manufacturing* 28(1): 75-86.
16. Ghomi Einollah Jafarnejad, Amir Masoud Rahmani, Nooruldeen Nasih Qader. 2019. "Cloud manufacturing: challenges, recent advances, open research issues, and future trends". *The International Journal of Advanced Manufacturing Technology* 102(9): 3613-3639.
17. Narwane Vaibhav S., Balkrishna E. Narkhede, Bhaskar B. Gardas, Rakesh D. Raut. 2019. "Cloud manufacturing issues and its adoption: past, present, and future". *International Journal of Management Concepts and Philosophy* 12(2): 168-199.
18. Issues and its adoption: past, present, and future. *International Journal of Management Concepts and Philosophy* 12(2): 168-199.
19. AbubakarA. D., Julian M. Bass, Ian Allison. 2014. "Cloud computing: Adoption issues for sub-saharan African SMEs". *The Electronic Journal of Information Systems in Developing Countries* 62(1): 1-17.
20. Huang Biqing, Chenghai Li, Chao Yin, Xinpei Zhao. 2013. "Cloud manufacturing service platform for small-and medium-sized enterprises". *The International Journal of Advanced Manufacturing Technology* 65(9): 1261-1272.
21. Zhang Zhenjie, Yuanming Zhang, Jiawei Lu, Fei Gao, Gang Xiao. 2020. "A novel complex manufacturing business process decomposition approach in cloud manufacturing". *Computers & Industrial Engineering* 144: 106442.
22. Abd Rahman, M.N., B. Medjahed, E. Orady, M.R. Muhamad, R. Abdullah, A.S.M. Jaya. 2018. "A review of cloud manufacturing: Issues and opportunities". *Journal of Advanced Manufacturing Technology (JAMT)* 12(1): 61-76.
23. Fisher Oliver, Nicholas Watson, Laura Porcu, Darren Bacon, Martin Rigley, Rachel L. Gomes. 2018. "Cloud manufacturing as a sustainable process manufacturing route". *Journal of manufacturing systems* 47: 53-68.
24. Jula Amin, Elankovan Sundararajan, Zalinda Othman. 2014. "Cloud computing service composition: A systematic literature review". *Expert systems with applications* 41(8): 3809-3824.
25. Adamson Göran, Lihui Wang, Magnus Holm, Philip Moore. 2017. "Cloud manufacturing–a critical review of recent development and future trends". *International Journal of Computer Integrated Manufacturing* 30(4-5): 347-380.
26. Xu X. "Cloud manufacturing: A new paradigm for manufacturing businesses". 2013. *Australian Journal of Multi-Disciplinary Engineering* 9(2): 105-116.
27. Wang Peng, Robert X. Gao, Zhaoyan Fan. 2015. "Cloud computing for cloud manufacturing: benefits and limitations". *Journal of Manufacturing Science and Engineering* 137(4).
28. Tao Fei, Lin Zhang, V. C. Venkatesh, Y. Luo, Ying Cheng. 2011. "Cloud manufacturing: a computing and service-oriented manufacturing model". *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* 225(10): 1969-1976.

29. Marston Sean, Zhi Li, Subhajyoti Bandyopadhyay, Juheng Zhang, Anand Ghalsasi. 2011. "Cloud computing—The business perspective". *Decision support systems* 51(1): 176-189.
30. Wang Xi Vincent, Xun W. Xu. "An interoperable solution for cloud manufacturing". 2013. *Robotics and computer-integrated manufacturing* 29(4): 232-247.
31. Agrawal Divyakant, Amr El Abbadi, Sudipto Das, Aaron J. Elmore. "Database scalability, elasticity, and autonomy in the cloud". 2011. In: *International Conference on Database Systems for Advanced Applications*: 2-15. Springer, Berlin, Heidelberg.
32. Zhao Chun, Xiao Luo, Lin Zhang. "Modeling of service agents for simulation in cloud manufacturing". 2020. *Robotics and Computer-Integrated Manufacturing* 64: 101910.
33. Wattal Suneel, Ajay Kumar. 2014. "Cloud computing-an emerging trend in information technology". In: *2014 international conference on issues and challenges in intelligent computing techniques (ICICT)*: 168-173. IEEE.
34. Zissis Dimitrios, Dimitrios Lekkas. 2012. "Addressing cloud computing security issues". *Future Generation computer systems* 28(3): 583-592.
35. Saaty Thomas L. 1994. "How to make a decision: the analytic hierarchy process". *Interfaces* 24(6): 19-43.
36. Hu Yanjuan, Feifan Zhu, Lin Zhang, Yongkui Lui, Zhanli Wang. 2019. "Scheduling of manufacturers based on chaos optimization algorithm in cloud manufacturing". *Robotics and Computer-Integrated Manufacturing* 58: 13-20.
37. Hu Yanjuan, Lizhe Wu, Chao Shi, Yilin Wang, Feifan Zhu. 2020. "Research on optimal decision-making of cloud manufacturing service provider based on grey correlation analysis and TOPSIS". *International Journal of Production Research* 58(3): 748-757.
38. Sevinc Ali, Şeyda Gür, Tamer Eren. 2018. "Analysis of the difficulties of SMEs in industry 4.0 applications by analytical hierarchy process and analytical network process". *Processes* 6(12): 264.
39. Mian Syed Hammad, Bashir Salah, Wadea Ameen, Khaja Moiduddin, Hisham Alkhalefah. 2020. "Adapting universities for sustainability education in industry 4.0: Channel of challenges and opportunities". *Sustainability* 12(15): 6100.
40. Dağdeviren, Metin. 2008. "Decision making in equipment selection: an integrated approach with AHP and PROMETHEE". *Journal of intelligent manufacturing* 19(4): 397-406.
41. Luthra Sunil, Sachin Kumar Mangla. 2018. "Evaluating challenges to Industry 4.0 initiatives for supply chain sustainability in emerging economies". *Process Safety and Environmental Protection* 117: 168-179.

APPENDIX-1
Survey questionnaire (to be filled by experts)

You are supposed to compare two opportunities at a time (that is, in pairs). The scores of comparisons are 1, 3, 5, 7, 9. Scores are assigned as:

Equal importance	1
Moderate importance	3
Strong importance	5
Very strong importance	7
Extreme importance	9

For example, if you are comparing opportunity (O_1) of a row with opportunity (O_2) of the column and assigned value **5**, it means that opportunity O_1 is of strong importance than opportunity O_2.

NOTE: No need to fill cells with value 1.

Opportunities in the espousal cloud manufacturing

		Opportunities in column i								
Opportunities in row i		0_1	0_2	0_3	0_4	0_5	0_6	0_7	0_8	0_9
	0_1	1								
	0_2		1							
	0_3			1						
	0_4				1					
	0_5					1				
	0_6						1			
	0_7							1		
	0_8								1	
	0_9									1

Table 1

No	Opportunities
O_1	Pay-per-use or Instant service
O_2	Scalability
O_3	Cost efficiency
O_4	Flexibility
O_5	Autonomy
O_6	Low-risk backup and recovery
O_7	Low startup cost
O_8	Location independence

Received 15.02.2022; accepted in revised form 30.03.2022

Scientific Journal of Silesian University of Technology. Series Transport is licensed under a Creative Commons Attribution 4.0 International License

Proceedings of the Seventh International Conference on Communication and Electronics Systems (ICCES 2022)
IEEE Xplore Part Number: CFP22AWO-ART; ISBN: 978-1-6654-9634-6

2022 7th International Conference on Communication and Electronics Systems (ICCES) | 978-1-6654-9634-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCES54183.2022.9835933

SentiGames- A Game Theoretic Approach To Sentiment Analysis

Minni Jain[1], Banavath Gayathri[2], Ritwik Ranjan[3], Rohit[4]
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
[1] minnijain@dtu.ac.in [2] banavathgayathri_2k18co107@dtu.ac.in
[3] ritwikranjan_2k18co296@dtu.ac.in [4] rohit_2k18co301@dtu.ac.in

***Abstract*—Sentiment Analysis makes use of several natural language processing methods in order to understand subjective knowledge from a given statement. This is an important issue that is presently being explored because of the multiple applications that may be derived from comments on social media platforms, answers from surveys, product assessments, or making data driven choices. Our work suggests an unique evolutionary game theoretic method to sentiment analysis. Each word in our method is represented as a node on a graph, and each sentiment-sense as a class. The algorithm distributes sentiment-sense to consistent classes depending on how similar each word is to the others, so that identical words are confined to comparable classes. The system's dynamics are modeled as a non-cooperative multiplayer game in which players are data points who pick their class affiliations and equilibria correlate to accurate data labeling. By merging many modern game theory methods, we hope to find the best feasible solution using game theory to optimize the accuracy of the specified NLP task, namely Sentiment Analysis.**

***Index Terms*—Sentiment Analysis, GloVe Word Embedding, Natural Language Processing, Sentiwordnet, Game Theory**

I. Introduction

Nowadays, we notice a lot of automatic actions taking place in our smart devices by simply uttering the command or reading our emails, among other things. These activities fall under the umbrella of Natural Language Processing, a branch of computer science. To perform NLP functions, a variety of models and algorithms are available, including Deep Learning models, such as LSTM, the Naive Bayes algorithm, and others. All of these systems, on the other hand, rely on manually labelled data, which can be time intensive when dealing with thousands of items. This is when unsupervised learning comes into play. After reading a few research publications, we concluded that Game Theory is the key to performing extremely accurate NLP jobs without a large dataset or manually annotated data. We may readily interpret different features of the tasks as players playing a game using modern game theory algorithms, and we can select the optimal response by choosing the player with the biggest payoff while playing their best strategies.To extract information from a statement, categorisation of feelings(sentiments) today makes use of a variety of treatment of natural speech strategies. Our study proposes a revolutionary, a game-theoretic analysis of evolution approach.

II. Related Work

In this part, we will first examine previous work in sentiment analysis before exploring significant applications of game theory in natural language processing.

A. Sentiment Analysis

Appraisal of Public Opinion Sentiment analysis is a well-known field that recognises and categorizes feeling, judgment, and expression in written content by humans. Prior polarity lexicons get added in a standard heuristic algorithm for sentiment analysis, where entries are labeled with their initial from out context polarity as humans perceive it using their learning outcomes. The focus of this research is to determine how similar word pairs are in terms of emotion. By comparing sentiment similarity between two words, you can determine how far apart they are in terms of their underlying attitude. To date, many ways to capture the semantic similarity between words have been developed, for example, the WordNet-based comparison method for semantic similarity metrics.

B. Sentiment Analysis Approaches

Sentiment Analysis is a division of NLP that seeks to extract attitudes, assessments, opinions, and emotions from the text. Sentiment Analysis has become an essential aspect of the analytics firms must undertake to understand how they have positioned in the market. Sentiment Analysis may approach in various ways; we'll look at a few of them.

1) Naive Bayes: The Naive Bayes classifier makes predictions based on prior knowledge of circumstances that may be associated using probabilities. Put another way; it calculates the outcome based on the conditional probability of each lexical characteristic Preprocessing is required for NLP machine learning to extract features from the original text. Vectorization is the process of converting text material into numerical forms so that algorithms can handle it. For Naive Bayes, we will create a simple DTM (document-term matrix) for model consumption. However, it can incorporate additional text length, named entities, etc.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:12:12 UTC from IEEE Xplore. Restrictions apply.

2) Deep Learning LSTM: Deep Learning allows for considerably more nuanced input to be absorbed. The LSTM model, is indeed a sort of RNN that would be used to manage sets of data. We employ this neural network design because we believe the order of characteristics (words) in a sentence is essential. Deep learning is resource intensive, and it has trouble exposed to high, terms of features (poor throughput and leisurely coherence). We must encode features extracted from the original text as dense vectors when using them in model training. One way is to convert each reader into a sequence of numbers, each of which corresponds to a lexical word [12]. We may move beyond by utilizing Word Embeddings to map words with comparable. The training data can find this in both positive and negative text—usage/meaning to similar actual number vectors (rather than an index).

C. Game Theory and NLP

In dynamic choices and logical frameworks, game theory delivers predictive power. Von Neumann and Morgenstern invented game theory as the basic notion that prompted the invention of the mathematical model used to simulate decision-making issues [6]. Smith and Price proposed an evolutionary game theory to improve classical game theory by eliminating subjective terms such as the player's hyper-rationality [9]. In order to enhance the effectiveness of information extraction, game theory has been employed as a decisional paradigm for content sorting. It replicates users' interactions with search engines as a game and then optimizes the system as a whole to boost interactive information retrieval performance. Game-theoretic approaches have also been applied to select the optimal conversational bot replies. Further, Tripodi also adopts a game-theoretic model for context meaning and word sense disambiguation, where a graph nodes represent each word to be disambiguated, edges reflect contextual interactions, and word senses are represented as strategy space classes [1].

D. Rumour Source Detection using Game Theory

Since people are involved on social networking sites which can also be a means to spread rumours on the internet. But it might also be possible that all these facts are not true which may include unsupported rumours. Therefore, this paper plays a crucial role in finding the prominent root of spreading rumours [14]. The most basic visualization is to classify the node and edges from most severe to least severe. They used five SV-based games to find the top 10 most prominent nodes on three network datasets and compared the results using basic centrality metrics, resulting in an SV-based method for better understanding marginal contribution and the true influence of each node on the whole network. They also inspected five various game circumstances, taking into analysis of various approaches to determine the most prominent nodes in a given dataset which assisted to earn an extensive and total understanding of the game-theoretical algorithm.

III. Preliminaries

Game theory basically is a branch of applied mathematics concerned with the study of games provides methods for studying scenarios in which two or more persons, referred to as players, make interdependent decisions. Because of this interdependence, while deciding on a technique, each participant must take into account the other's decisions or plans. A game solution describes the ideal options of the players, who may have similar, opposing, or eclectic interests, the outcomes that may result from these actions. Understanding the position on the basis of the behaviors given to users including associated paybacks as a consequence of occupations, as well as utilizing alternative equilibrium concepts to make diagnostic or didactic predictions, are essential components of doing a noncooperative game recaps [4]. Estimating all parties' costs and benefits of each possible set of options are the first step in determining payoffs. Estimating gains for some people (like yourself) may be easier than for others in many scenarios, and it's likely that some individuals aren't evaluating it in the same manner you are. A plan of action w_i is player k's best reaction of player to a profile of approaches $w_{-k} \epsilon w_{-k}$ for either opponents, also known as the best response if:

$$v_k(w_k, w_{-k}) \geq v_k(w_k^*, w_{-k}), \text{ for each } w_k^*. \quad (1)$$

A player's best answer to the characteristics of other players' strategies is a strict perfect reply if it is the only best response. A summary of methods $w \epsilon W$ is a strategy in its purest form. if w_k is the best solution to w_k for each k then We've arrived at a Nash equilibrium.

$$v_k(w_k^*, w_{-k}) \leq v_k(w_k, w_{-k}), \text{ for each k and } w_k^*. \quad (2)$$

A. Evolutionary Game Theory

Evolutionary game theory is a development of classical game theory proposed by John Maynard Smith and George Price that removes the players hyperrationality constraint. [9] If we consider situations from the real world we can observe that strategies are affected by social norms and heuristics also. It introduces a novel concept of inductive learning in which the players repeatedly play games with their neighbors and their approach based on prior games' payoffs The strategy space of a player is defined as a mixed strategy profile xi, is contained inside a mixed strategy space assessed by the cartesian product in this example:

$$\Theta = \times_{i \in I} \Delta_i \quad (3)$$

What distinguishes evolutionary game theory from classical game theory is that the player receives a final payout after playing with all of the other players, which is calculated by adding the payoffs of all the single games the player has played, as follows:

$$u_i(e_i^h) = \sum_{j=1}^{n} (A_{ij} x_j)_h \quad (4)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:12:12 UTC from IEEE Xplore. Restrictions apply.

Whereas the average payoff will be computed as:

$$u_i(x) = \sum_{j=1}^{n} x_i^T A_{ij} x_j \quad (5)$$

where n denotes the total number of players with whom the player has played games, and Aij denotes the reward matrix between players I and j. Another fundamental property of game theory is that the player plays the game repeatedly and at each chance of playing, they have the privilege of updating their strategy space according to the payoffs they received in the previous games in order to increase the probability of strategy from their mixed strategy which increases the payoff till the point that the strategy space can not be further updated as the maximum payoff that the player can receive is attained and any further increase is impossible.

$$x' = [u(eh, x) - u(x, x)].xh \forall h \in S \quad (6)$$

equation 6 can be used in a dynamic system on the basis of frequency based selection. It works on the assumption that growth rate of each strategy is directly proportional to its fitness, which is later defined by summing the payoffs gained during the games due to which several strategies can exist side by side with different rates.

Word Embedding

Glove: We used pre-trained GloVe to train our game theory model to investigate whether adding the IMDB corpus to global statistics enhances game-theoretic model performance. GloVe, as an unsupervised trained model, assists in the mapping of words into a significant space in which semantic similarity is proportional to word distance [7] . There are two parts to the dataset: training and testing.

SentiWordNet: SentiWordNet is a viewpoint dictionary generated from the WordNet repository, with statistical results reflecting favorable and unfavorable sentiment polarity on every term. However, one database is SentiWordNet, which comprises viewpoint statistics on phrases collected again from the WordNet repository and made publicly available for study purposes [10]. A wordNet is a dictionary consisting of English keywords that have already been categorized as analogs as sequences. Although WordNet can be referred to as a vocabulary, it is thought to be better linguistically precise because it contains alternatives to terms in particular cases [11]. The database should be preprocessed, which includes the elimination of stop-words and diacritics. The sentences can be kept in Py databases to make them easier to handle. It is critical to determine the Words and Expressions for each word in the dictionaries whenever utilizing SentiWordNet. Expressions include verbs, nouns, adverbs, and adjectives. Conjunction of Prepositions Interjection with a pronoun When reviewing the sentiments of a sentence, the verb, noun, and adjectives are the most usually employed.

IV. Proposed Work

Dataset Description- However, we have trained our model in a semi-supervised way, but for that, too, we need some quantity of data to prepare in order to produce accurate results. Observing the popularity of sentiments in various movie reviews, the dataset is taken here was the aclImdb dataset.A total of 25,000 IMDB reviews were used to create the following dataset. In the sample, there were an equal number of strong and poor reviews. If the mark is less over 4, it is considered adverse; if the rating is greater than 7, the rating is deemed favorable.There were equitably distributed 25,000 assessments throughout the simulation and experimental categories.

	review	sentiment
0	One of the other reviewers has mentioned that ...	positive
1	A wonderful little production. The...	positive
2	I thought this was a wonderful way to spend ti...	positive
3	Basically there's a family where a little boy ...	negative
4	Petter Mattei's "Love in the Time of Money" is...	positive

Figure 1: IMDB Data

Splitting the dataset: We separated the collection into the training set and testing in what appeared to be a 4:1 proportion. Due to vast size of the dataset and in order to conserve computing power and time, the holdout approach is utilized. We'll utilize both sets of data to develop the model once we've segregated the dataset into training and testing sets.

Preprocessing- From the Dataset consisting of 25000 movie reviews, 80% are used for the model's training. There may be some unwanted data in the Dataset that has to be eliminated because it is uneven, as seen in figure 1.

To eliminate inconsistencies in our data, we must undertake pre-processing operations to ensure that it is excellent input for our model training.

- Stopwords Removal using NLTK- To get relevant data from our dataset, we need to get rid of Stop words first. Stop words such as the, a, and an do not add much to the comprehension of the text. We needed a consistent structure after deleting stopwords, thus all lowercase and uppercase words were changed to lowercase to provide suitable semantics.We also get rid of characters that aren't useful in the public domain, or aren't relevant.

- Tokenization- It was formerly used to break down longer text sequences into smaller chunks known as tokens. We Executed many pre-processing actions on the content once tokenization got accomplished.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:12:12 UTC from IEEE Xplore. Restrictions apply.

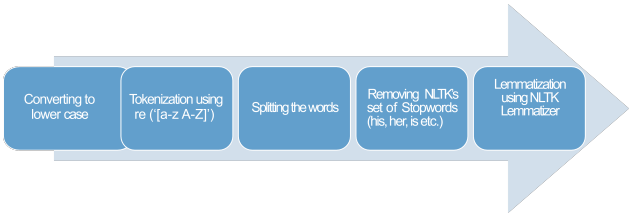


Figure 2: Data Preprocessing Technique

	review	sentiment
0	one reviewer mentioned watching oz episode you...	1
1	wonderful little production filming technique ...	1
2	thought wonderful way spend time hot summer we...	1
3	basically there family little boy jake think t...	0
4	petter matteis love time money visually stunni...	1

Figure 3: Data After Preprocessing

So, after applying all the pre-processing tasks which are mentioned in figure 2, we obtain clean data as shown in figure 3, which can be used for our model training.

A. Working Of Our Model

Our work suggests an unique evolutionary game-theoretic method to sentiment analysis. Each word in our method is represented as a node on a graph, and each sentiment-sense as a class. The algorithm distributes sentiment-sense to consistent classes depending on how similar each word is to the others, so that identical words are confined to comparable classes. The system's dynamics are modeled as a non-cooperative multiplayer game in which players are data points who pick their class affiliations and equilibria correlate to accurate data.

The applicability of the game theory is such that the ordinary person's decisions always seem to have an opportunity to choose from in a manner that produces an even less best possible outcome for both the individuals as a collective, and the utilization of behavioral economics can support players reach an effective decision because once actually faced by unbiased and trying to compete performers in a crucial first step.

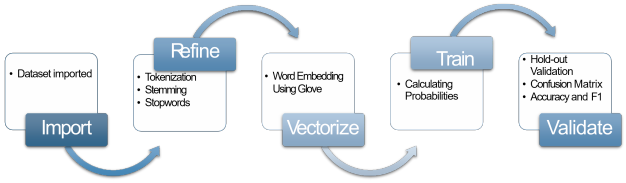


Figure 4: Model Architecture of our evolutionary stable method

In this research, we would like to propose novel ideas around the prediction of the sentiment of a sentence without any prior data evidence using the Game Theoretical Machine Learning model built using Senti WordNet Lexicon. We will further embed the data using modern word lexicons like Glove, WordNet, Bert, etc. First, we observed that the evolutionary game theory model replicates the development of a language over time. Hence, we can deduce that computational linguistics can be easily linked to the evolutionary game theory.We can also see that the feeling of a statement is driven by all of the words in the sentence. Each word influences the sentiment expressed by other words, hence driving the overall sentiment of the sentence towards one specific sentiment. Our model selected different words as players and assigned payoffs according to their sentiments. We also assigned the similarity and distance measures between the words so the overall sentiment of the statement will be determined by its various senses in the strategic space.

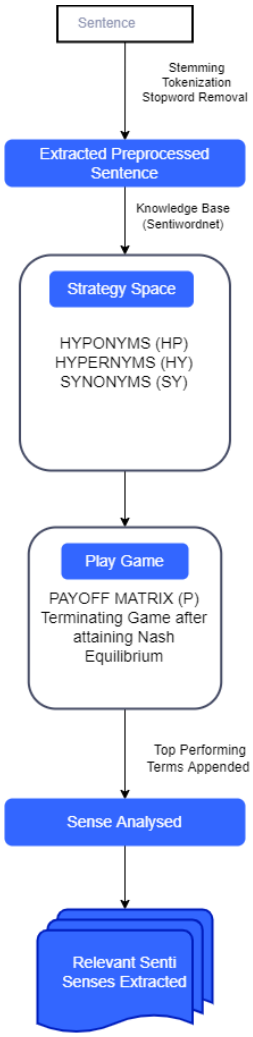


Figure 5: Framework for proposed methodology

Moving forward towards the framework for our proposed methodology in figure 5, the sentence will first be running through our preprocessor, which will remove stopwords, tokenize, and embed the words. We next build the strategy space using the provided word's hyponyms, hypernyms, and synonyms. We also construct the payoff matrix by combining

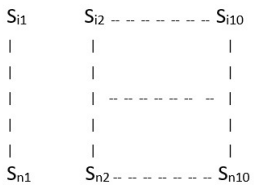
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:12:12 UTC from IEEE Xplore. Restrictions apply.

the similarity index and the word's relative distance. The extracted strategy space and payoff matrix are then used to play games until the Nash Equilibrium is reached. By merging the sense assessments, the ultimate sentiment is formed.

B. The Strategy Space

We'll talk about how we built strategy space in our game-theoretic model in this section. Now we must define the actions that each player can take, also known as the player's strategy space. The players in our game are words, and their strategies should be designed to maximize their output based on the payoff matrix. The payoff matrix will be discussed in the next section. The different strategies in our case will be nothing more than the various senti-senses of the word, up to a maximum of 10 senti-senses. As a result, each row will correspond to one sentence word, and each value in that row will represent the probability that a particular senti-sense will be the actual representation of the word.

So, if there are n words and each of them has 10 meanings, the strategy space will look like this:



Each row in the above matrix relates to a player's strategy space and travels up to n rows, while each column corresponds to senti-sense and moves up to 10 columns.

C. The Payoff Matrix

The Payoff Matrix in game theory is defined as a matrix containing the payoffs or the awards that the players will get according to the strategy they will play. Now in the referenced paper of Word Sense disambiguation [1], we can see that the payoff matrix is formed for bigrams only, but in another research paper that uses Naive Bayes [13], it is observed that sentiment analysis unigrams work better than bigrams, and hence we only incorporated the word similarly that is obtained by finding the cosine distance between the two-word vectors that originated from Glove Word Embedding and hence the payoff matrix will give more payoffs to similar words rather than not so identical words. Hence our overall prediction of the sentence will lie on the assumption that similar words in a sentence depict identical sentiments. The vector size of the payoff matrix will be the same as of strategy space as each strategy will correspond to a particular payoff.

The demerits which we've faced are that the Game theory presupposes that each firm is aware of the other's strategy compared to its own and can develop a pay-off matrix for a plausible conclusion. It is an entirely ridiculous and unworkable premise. The game theory implies that both nodes can be wise individuals. The numerous techniques used by one competitor versus another contribute to that unending and very impractical reasoning process. In its generalization to plutocracy, game theory is impractical since it assumes a constant-sum game. Similarly, the merits are According to Game Theory, plutocracy rates are usually well-enforced. Numerous economic concerns incorporating hazard and scientific relationships exist. The testable computation of games can help you cope with problems. In theory, a player is often a mere individual or a company throughout the physical world that would have to make decisions despite constrained capacity. A player is often a very individual or a company throughout the physical world that would have to make decisions despite constrained capacity. Because the pay-off quantities are so important, a player's comprehensive understanding of such pay-off matrix entails flawless estimates of any and all aspects influencing the execution of various tactics.

V. Results And Analysis

From the dataset, we've gathered 25000 reviews. The model was split into two datasets: train and test constitute about 80% and 20% of the statistics, accordingly.

We utilized accuracy and confusion matrix measurements to determine the correctness of the model we constructed.After charting the confusion matrix between positive and negative sentiments, we calculated the outcomes.

Using a variety of natural language processing techniques and algorithms, we calculated the findings for both supervised and unsupervised learning.

The following are the outcomes of the set:

Algorithm	NLP Technique	Overall Accuracy
Game Theory (Unsupervised)	Word Embedding Without GloVe	0.66
Game Theory (Unsupervised)	Word Embedding With GloVe	0.62
Multinomial Naive Bayes (Supervised)	TF-IDF Vectorisation-Unigram	0.86
Multinomial Naive Bayes (Supervised)	Count Vectorisation-Unigram	0.85
Logistic Regression (Supervised)	TF-IDF Vectorisation-Unigram	0.89
Logistic Regression (Supervised)	Count Vectorisation-Unigram	0.88

Table 1: Comparison of scores

Basically, accuracy is computed using the ratio of total correct predictions to the total number of predictions, i.e., the sum of true positives and true negatives divided by the total number of tests. The accuracy of models that are supervised is a bit more higher due to the presence of corpora knowledge. In contrast, game theory is unsupervised, and no prior knowledge

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:12:12 UTC from IEEE Xplore. Restrictions apply.

of the dataset is present with the model. Hence the accuracy is as high as other supervised models but higher than most of the unsupervised models currently present.

As seen in the results, Game Theory Model Without Word Embedding proved to be much more accurate than the one with GloVe Word Embedding, with an accuracy score of 66% for Unsupervised Learning whereas for supervised learning Logistic Regression Model with TF-IDF Vectorisation proved to be more accurate, with an accuracy score of 89%.

VI. Conclusion And Future Work

After quite a series of tests on our research, we can conclude that Game Theory Model without GloVe Word Embedding proven to deliver greater accuracy scores for unsupervised learning. Our approach to sentiment analysis is based on an unsupervised game-theoretic methodology, which eliminates the requirement for a large human-annotated dataset.We noticed that each word has an impact on the sentiments expressed by other words, causing the overall sentiment of the phrase to gravitate toward a single sentiment.To bias the weight and payout matrices, additional characteristics may be included. For our input sentences, a generic word embedding strategy like BERT with the semantic representation may provide superior results.We expect that our game-theoretic approach to this problem will lead to additional innovative research in this area.

References

[1] Rocco Tripodi, Marcello Pelillo. A Game-Theoretic Approach to Word Sense Disambiguation, Computational Linguistics (2017). 2017 Association for Computational Linguistics Published under a Creative Common Attribution NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) license.

[2] J.Nash. Non-cooperative games. Annals of mathematics, pages 286-295, 1951.

[3] Jeffrey Pennington, Richard Socher, Christopher D. Manning GloVe: Global Vectors for Word Representation. https://www.aclweb.org/anthology/D14-1162 (2015)

[4] Rocco Tripodi, Roberto Navigli. Game Theory Meets Embeddings: A Unified Framework for Word Sense Disambiguation. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015).

[5] Rocco Tripodi Marcello Pelillo. WSD-games: a Game-Theoretic Algorithm for Unsupervised Word Sense Disambiguation. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015).

[6] J. Von Neumann and O. Morgenstern. Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition). Princeton University Press, 1944.

[7] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation.

[8] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. CoRR, abs/1301.3781, 2013.

[9] Smith J M, Price G R(1973) "The logic of animal conflict." Nature. vol.15. pp. 246.

[10] Esuli, Andrea Sebastiani, Fabrizio. (2006). SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining.

[11] Miller G A (1995) WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41.

[12] Minaee, Shervin Azimi, Elham Abdolrashidi, AmirAli. (2019). Deep-Sentiment: Sentiment Analysis Using Ensemble of CNN and Bi-LSTM Models.

[13] R. Ranjan and P. Yadav, "Emoji Prediction using LSTM and Naive Bayes," TENCON 2021 - 2021 IEEE Region 10 Conference (TENCON), 2021, pp. 284-288, doi: 10.1109/TENCON54134.2021.9707344.

[14] Jain, Minni Jaswani, Aman Mehra, Ankita Mudgal, Laqshay. (2020). Rumour Source Detection Using Game Theory. International Journal of Interactive Multimedia and Artificial Intelligence. 6. 49. 10.9781/ijimai.2020.10.003.

[15] Jain, M., Suvarna, A. Jain, A. An evolutionary game theory based approach for query expansion. Multimed Tools Appl 81, 1971–1995 (2022). https://doi.org/10.1007/s11042-021-11297-x

[16] Pandian, A. Pasumpon. "Performance Evaluation and Comparison using Deep Learning Techniques in Sentiment Analysis." Journal of Soft Computing Paradigm (JSCP) 3, no. 02 (2021): 123-134.

[17] Sungheetha, Akey, and Rajesh Sharma. "Transcapsule model for sentiment classification." Journal of Artificial Intelligence 2, no. 03 (2020): 163-169.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on August 22,2022 at 11:12:12 UTC from IEEE Xplore. Restrictions apply.

Soft Computing
https://doi.org/10.1007/s00500-022-07397-x

FOCUS

SIRA: a model for propagation and rumor control with epidemic spreading and immunization for healthcare 5.0

Akshi Kumar[1] · Nipun Aggarwal[2] · Sanjay Kumar[2]

Accepted: 14 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Healthcare social networks have a significant role in providing connected and personalized healthcare environment with real-time capabilities. However, building resilient, robust and technology-driven healthcare 5.0 has its own barriers. Especially with social media's high susceptibility to rumors and fake news, these networks can harm the society. Many researchers have been investigating the process of information diffusion, and it has been one of the most intriguing issues in network analysis. Modeling rumor propagation is one of the prominent researched topics in recent years. Traditional models assume that rumor propagation happens only in one direction, where only supporters are supposed to be active, whereas, in a real-life situation, both supporters and deniers of the information operate simultaneously. In this paper, we introduce a model for the recovery of nodes in a setting where rumor propagation and rumor control happen simultaneously. We propose the Susceptible-Infected-Recovered-Anti-spreader model based on the notion of spreading of epidemics and also its applications to modeling the propagation of rumors and control of rumor. Our model assumes people have multiple forms of reactions to rumor, either posting it, deleting it or announcing the rumor as fake. This paper also suggests how the model can act as a simulation method to compare two node centrality algorithms where spreaders chosen from one centrality algorithm try to spread the rumor, and the anti-spreaders chosen from other centrality try to dispel the rumor and vice versa. We simulate the proposed algorithm on different weighted and unweighted real-world network datasets and establish that the experimental results agrees with the proposed model.

Keywords Complex network · Healthcare social networks · Information diffusion · Node centrality · Rumor control · Rumor propagation

1 Introduction

The 5.0 revolution in healthcare is eventually digitally transforming society, plummeting the overload on emergency and critical care, and letting people to reinforce their resilience and health with a faster response time. The use of AI, the Internet of Medical Things (IoMT) and social media is offering number of advantages over traditional analytic and clinical decision-making techniques. In this view, many healthcare social networks (HSN) are evolving with the intent to augment patient care and awareness enticing the interest of clinical professionals and of the whole healthcare industry. Technological developments and digital platforms foster information diffusion possibilities to keep people safe, informed and connected. However, the same tools also expose possible perils of misinformation for patients. As the world combats with the outrageous and perilous coronavirus, the pandemic has demonstrated how the spread of misinformation, amplified on social media and other digital platforms, is proving to be as much a threat jeopardizing measures to control the pandemic.

Typically, rumors are information pieces that are unreliable, incorrect and spread on the network due to limited

Communicated by Deepak kumar Jain.

✉ Sanjay Kumar
sanjay.kumar@dtu.ac.in

Akshi Kumar
akshi.kumar@mmu.ac.uk

Nipun Aggarwal
nipundtu@gmail.com

[1] Department of Computing and Mathematics, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, United Kingdom

[2] Department of Computer Science and Engineering, Delhi Technological University, New Delhi 110042, India

Published online: 12 August 2022

knowledge of the recipient nodes or wrong intentions (Qiu et al. 2021; Spohr 2017; Sivasankari and Vadivu 2021). Rumors are the unverified spread of societal concerns, issues by different methods, and create misunderstandings between individuals. Rumors on the Internet spread quicker and have a more extensive scope of an impact than word-of-mouth rumors. These rumors spread by word of mouth over a small area and in vast geographical regions like a forest fire on social networks (Pfeffer et al. 2013), generally with malicious intent. Rumors alter and shape public opinion adversely affecting human behavior, causing multiple confusions (Lai et al. 2020) and instability. Many researchers have been investigating the process of information diffusion, and it has been one of the most intriguing issues in network analysis. Modeling rumor propagation is one of the prominent researched topics in recent years. The SIR (Harko et al. 2014) model is a traditional model that simulates information diffusion due to similarity in the field of epidemiology and social networks. It has been utilized by numerous researchers to model the process of information diffusion in social networks. A susceptible individual spreads data when influenced by a spreader and turns into a spreader. After some time, the spreader quits spreading the data and turns into a recovered node. The transformation parameters β and γ in the process are network characteristics-based probabilities, and the state transformation of all nodes is equivalent. While the SIR model serves as an excellent benchmark for information propagation, it does not generalize well to rumor propagation. A notable difference arises because individuals may also be immunizing other nodes regarding the rumor and preventing them from getting infected. This phenomenon is very different from actual epidemics modeling because, in an epidemiologic setting, it is unlikely that a large portion of people who recover from the epidemic, make other people immune as well. However, in a social network, when a person propagated fake news intentionally or due to a lack of knowledge and after getting the truth, other persons can disseminate the fact about the misinformation. This makes their susceptible neighbors, directly recovered and immune to the rumor. We consider the above recovery aspects of information propagation in our paper and propose the Susceptible-Infected-Recovered-Anti-spreader (*SIRA*) model based on the notion of spreading of epidemics and its applications to modeling the propagation of rumors and control of rumor. Our model assumes people have multiple forms of reactions to rumor, either posting it, deleting it or announcing the rumor as fake. The contribution of our papers are as follows:

- We establish mathematical model for the rumor propagation, rumor control and analyzing its equilibrium parameters and reproduction rate.
- The proposed model introduces the notion of anti-spreaders that immune its susceptible neighbors from the rumors to minimize the effect of rumors in the network.
- We consider the case where the information is not a rumor, and the proposed model simplifies to information propagation model for the influence maximization.
- We discuss the novel approach of directly comparing two different node centrality algorithms for the selection of spreaders and anti-spreaders. The spreaders, chosen from one centrality algorithm, try to spread the rumor, and the others try to dispel the rumor and vice versa.

The paper organization is as follows. In Sect. 2, we present some traditional models used to study rumor propagation and clarify the enhancements that our model proposes. In Sect. 3, we describe the proposed SIRA model of rumor propagation and rumor control with mathematical modeling using the system of differential equations. This section also explores the analysis of the proposed model on homogeneous networks, calculation of reproduction number and how our model can be simplified to a general information diffusion model as the SEI model followed by the simulation of the model on a toy network. In Sect. 4, we discuss the different weighted and unweighted real-world datasets used in this work, and we prove the mathematical integrity of our model for information propagation, rumor propagation, rumor control and node centrality algorithm performance comparison on various datasets used. Finally, Sect. 5 concludes the paper.

2 Related work

Social networks are complex networks with non-trivial topology and intricate interactions between users. The effective analysis of complex systems is essential to model and comprehend the method of rumor propagation. The spreading of rumors on online social networks is very similar to epidemic spreading. Researchers from various disciplines have extensively studied models of rumor propagation in complex networks. The majority of the conventional models for investigating information diffusion depend on elementary models of epidemics, for example SI (Susceptible-Infected), SIS (Susceptible-Infected-Susceptible) and SIR (Susceptible-Infected-Recovered) (Britton 2010; Kumar et al. 2020).

Tripathy et al. Tripathy et al. (2010) claimed that the effect of rumor can be combated by injecting messages called anti-rumors, just like the rumor process. They studied two metrics, namely the belief time, which is the duration for which a person believes the rumor to be true, and point of decline, which is a point after which the anti-rumor process dominates the rumor process. Huang and Jin Huang and Jin (2011) modified the SIR model to prevent the rumor propagation by devising immunization strategies, namely the random immunization

and the targeted immunization, to the rumor model on small-world networks. Wan et al. Wan et al. (2015) introduced the SIERsEs model, which is an extension of the SIR spreading model with the addition of the elimination of the rumor process. Zhu and Ma argue that certain people may not believe in the rumor and do not spread it, a different state called a hesitant state added to the traditional SIR model and proposed SHIR model (Zhu and Ma 2018). This model also takes care of the varying degree of nodes, which commonly happens in online social networks. Indu and Thampi argue that the spread of rumors in social networks is similar to the spread of fire in a forest and identified various components to model the spread of rumors in online social networks (Indu and Thampi 2019). Wang et al. Wang et al. (2019) explored the propagation of rumor using an improved energy model applicable to multilayer social networks. The model also introduces negative energy to simulate the mitigation process of rumors. Zhu et al. Zhu et al. (2020) recently proposed a SIR like an epidemic rumor propagation model by employing forced silence function, time delay and network topology. Recently, Yang et al. Yang et al. (2019) have developed an ILSR rumor propagation model based on the notion of the degree of different nodes in networks. They introduced a new type of users known as lurkers users who have heard the rumors, but temporarily not spreading rumors in the system. Kaur et al. Kaur et al. (2020) proposed a fake news detection system using multi-level voting model by exploiting various feature extraction techniques and machine learning classifiers. Qui et al. Qiu et al. (2021) introduced the SIR rumor-spreading model with an influence mechanism in social networks, named SIR-IM, by integrating the number of current spreaders into the spreading probability and time function.

3 Proposed methodology

3.1 *SIRA* model description

We propose the Susceptible-Infected-Recovered-Anti-Spreader ($SIRA$) model for the recovery of nodes in a setting where rumor propagation and rumor control happen simultaneously. The entire population is divided into four categories: Susceptible nodes (S) that are unaware of the rumor, Infected nodes (I) that are actively spreading the rumor, Recovered nodes (R) that are not spreading the rumor and anti-spreader nodes (A) that are dispelling the rumor. We consider people have multiple forms of reactions to rumor, either posting it, deleting it or announcing the rumor as fake. We introduce the notion of anti-spreader nodes, such nodes that get to know about the irrelevance of the information, and the piece of the information is a rumor, and tend to dispel the rumor by announcing the truth. Hence, the anti-spreaders nodes are immunizing their susceptible neighbor's nodes from false information and play a crucial role in controlling the rumor.

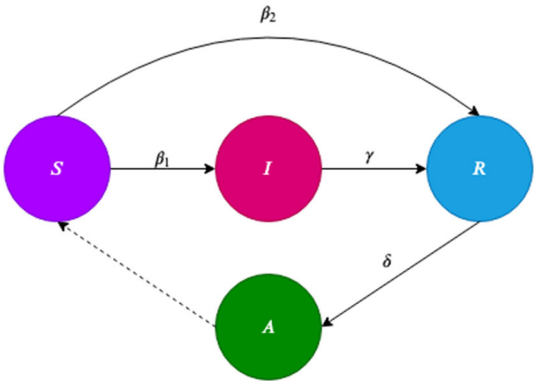


Fig. 1 $SIRA$ Model

As evident from Fig. 1, solid lines represent direct conversion of a node from one state to another, for examples $S- > A$, $I- > R$. However, the dotted line corresponds to indirect conversion from one state to other. For example, the conversion from A to S signifies the indirect effect of anti-spreaders on their susceptible neighbors. This makes susceptible neighbors transition to the recovered state with the rate β_2 and becomes immune to the fake news. The steps of the algorithm are as follows:

- Node population starts from certain infected nodes and certain anti-spreaders nodes that spread the rumor or dispel it, and all other nodes are susceptible
- Infected nodes affect their susceptible neighbors and convert them into infected members that propagate the rumor. This conversion happens with a probability β_1
- Probability that the infected members recover is at a rate of γ
- Under these conditions recovered nodes can now become anti-spreaders at the rate of δ
- Nodes that become anti-spreaders affect β_2 fraction of their susceptible neighbors, converting them directly into recovered state and immune to the rumor. Due to the influence of anti-spreaders nodes, some of their susceptible neighbors can directly go to a recovered state, and they will not be affected by the rumor.

Also, other than the rate of conversion β_1 from $S- > I$, the nodes in susceptible state can only transit to an infected state, if they are direct neighbors of an infected node. Similarly, other than the rate of conversion β_2 from $S- > R$, the susceptible nodes can directly transition into recovered nodes if they are direct neighbors of anti-spreader nodes.

At the end of simulation when there are no infected nodes, i.e., $I(t) = 0$ and equilibrium is reached then some nodes can still remain in susceptible state other then anti-spreaders.

It is not like that the final state is anti-spreader, and all the nodes should be in this state at the end; rather, it depends on the various conversion rates considered. Also, other than the rate of conversion β_1 from $S- > I$, the nodes in susceptible state can only transit to an infected state, if they are direct neighbors of an infected node. Similarly, other than the rate of conversion β_2 from $S- > R$, the susceptible nodes can directly transition into recovered nodes if they are direct neighbors of anti-spreader nodes.

3.2 System of differential equations

The population is divided into the 4 states Susceptible S, anti-spreaders A, Infected I and Recovered R and at any time t their sum is 1 as given in Eq. (1).

$$S(t) + A(t) + I(t) + R(t) = 1 \tag{1}$$

As evident from Fig. 1, some recovered nodes may become anti-spreaders at the rate δ as per Eq. (2). Anti-spreaders are those nodes who know the ingenuity of the information. The node in the anti-spreader(A) state continues to be in the same state throughout.

$$\frac{d}{dt}A(t) = \delta R(t) \tag{2}$$

Recovered nodes increase through multiple methods, either through the direct recovery of susceptible nodes due to influence of anti-spreaders (edge between S to R in Fig. 1) or by direct conversion from infected to recovered (edge between I to R in Fig. 1). Some recovered nodes may also become anti-spreader at the rate δ. The change in recovered nodes is presented in Eq. (3).

$$\frac{d}{dt}R(t) = -\delta R(t) + \gamma I(t) + \beta_2 * k * S(t) * A(t) \tag{3}$$

where k is the average degree of the nodes in the network. Change in infected nodes, Eq 4, arises from the conversion of susceptible nodes to infected nodes with the rate β_1 and infected nodes go to recovered state with the rate γ.

$$\frac{d}{dt}I(t) = -\gamma I(t) + \beta_1 * k * S(t) * I(t) \tag{4}$$

$$\frac{d}{dt}S(t) = -\beta_1 * S(t) * I(t) * k - \beta_2 * S(t) * A(t) * k \tag{5}$$

Initially, all the nodes are assumed susceptible except few seed nodes, and slowly, they decay into infected state with rate β_1 and recovered state with rate β_2 as given in Eq. (5). Now let us solve these equations by substituting $R(t)$ with $1 - S(t) - A(t) - I(t)$ from Eq. (3), and hence, getting the system given by 6, 7, 8.

$$\frac{d}{dt}A(t) = \delta(1 - S(t) - A(t) - I(t)) \tag{6}$$

$$\frac{d}{dt}I(t) = -\gamma I(t) + \beta_1 * k * S(t) * I(t) \tag{7}$$

$$\frac{d}{dt}S(t) = -\beta_1 * S(t) * I(t) * k - \beta_2 * S(t) * A(t) * k \tag{8}$$

3.3 Calculation of reproduction number

The basic reproduction number (R_0) is an important parameter in epidemics which measures the propagation potential of contagious diseases. The reproduction number (R_0) is an epidemiologic measure used to characterize the contagiousness or transmissibility of infectious agents. We have applied a similar concept for the rumor propagation. Since for successful reproduction of the rumor, we require that initially $\frac{d}{dt}I(t) > 0$, this produces Eq. (9)

$$-\gamma I(t) + \beta_1 * k * S(t) * I(t) > 0 \tag{9}$$

Thus, giving us Eq. (10),

$$R_t = (\beta_1 * k * S(t)/\gamma) > 1 \tag{10}$$

$R_t > 1$ means that the rumor is active in the system. For the initial spread of rumor, $R_0 > 1$, and since $S(0) \cong 1$, we get Eq. (11)

$$R_0 = (\beta_1 * k/\gamma) > 1 \tag{11}$$

Now, when $R_t < 1$, there is no rumor in the system. When $R_t < 1$ then $\frac{d}{dt}I(t) = 0$. This implies there is no infected node in the system, and hence, there is no rumor in the system.

3.4 Model analysis in a homogeneous network

We analyze the proposed $SIRA$ model on a homogeneous network and prove the validity of the model at the stable equilibrium points. An equilibrium point of a dynamic mathematical framework produced by an independent system of ordinary differential equation (ODEs) is a solution point where the count of participating entities does not change with time, and hence, their rate of change becomes 0 (Roussel 2005). In the case of epidemics, the disease-free equilibrium is asymptotically stable, and the infection asymptotically vanishes. Similarly, in the case of information propagation, rumor-free stable equilibrium is stable, and information is no longer being spread, and the network has no active rumor propagating nodes. We prove the stability of our differential equation at the equilibrium condition given by Eq. (12)

$$\frac{d}{dt}S(t) = 0, \frac{d}{dt}A(t) = 0, \frac{d}{dt}I(t) = 0, \frac{d}{dt}R(t) = 0 \tag{12}$$

The Jacobian method (Izhikevich 2007), as given by Eq. (13), is a popular iterative technique used for determining the solutions for the system of linear equations, which is diagonally dominant.

$$J_E = \begin{bmatrix} -\beta_1 * I * k - \beta_2 * A * k & -\beta_2 * S * k & -\beta_1 * S * k & 0 \\ 0 & 0 & 0 & \delta \\ \beta_1 * k * I & 0 & -\gamma & 0 \\ \beta_2 * k * A & \beta_2 * S * k & \gamma & -\delta \end{bmatrix} \tag{13}$$

The equilibrium points exist if I(t) = 0, and under this condition, we have Eq. (14)

$$E_i = (S, A, 0, R) \tag{14}$$

The Jacobian matrix is given by Eq. (15):

$$J_E = \begin{bmatrix} -\beta_2 * A * k & -\beta_2 * S * k & -\beta_1 * S * k & 0 \\ 0 & 0 & 0 & \delta \\ 0 & 0 & -\gamma & 0 \\ \beta_2 * k * A & \beta_2 * S * k & \gamma & -\delta \end{bmatrix} \tag{15}$$

We will calculate the eigenvalues of J_E as given in Eq. (16):

$$J_E = \begin{vmatrix} -\beta_2 * A * k - \lambda & -\beta_2 * S * k & -\beta_1 * S * k & 0 \\ 0 & -\lambda & 0 & \delta \\ 0 & 0 & -\gamma - \lambda & 0 \\ \beta_2 * k * A & \beta_2 * S * k & \gamma & -\delta - \lambda \end{vmatrix} \tag{16}$$

Expanding on the third row, we arrive at the following characteristic equation in Eq. (17):

$$(\gamma + \lambda)(\lambda)(\lambda^2 + (\delta - \beta_2 kA)\lambda - \beta_2 k\delta(A + S)) \tag{17}$$

As evident, $(-\gamma, 0)$ are eigenvalues and also the solutions of $\lambda^2 + (\delta - \beta_2 kA)\lambda - \beta_2 k\delta(A + S)$ which we find and analyze as follows using the Quadratic Formula:

$$\begin{cases} (\lambda^2 + (\delta - \beta_2 kA)\lambda - \beta_2 k\delta(A + S)) \\ D = ((\delta - \beta_2 kA)^2 + 4 * \beta_2 k\delta(A + S)) \\ \therefore \delta^2 + \beta_2^2 k^2 A^2 - 2\delta\beta_2 kA + 4\beta_2\delta kA + 4\beta_2\delta kS \\ \cong (\delta + \beta_2 kA)^2 + 4 * \beta_2 k\delta S \end{cases} \tag{18}$$

Here, since $(\delta + \beta_2 kA)^2 \geq 0 \because (x^2 > 0, \forall x \in R)$ and also $S, k, \beta_2, \delta \geq 0$, we have Eq. (19):

$$D \geq 0 \tag{19}$$

Thus, the eigenvalues are given by Eq. (20):

$$(0, -\gamma, (-(\delta - \beta_2 kA) - \sqrt{D})/2, (-(\delta - \beta_2 kA) + \sqrt{D})/2) \tag{20}$$

Now, for a stable equilibrium, real eigenvalues must be ≤ 0, and since $\gamma, \delta, \beta_2 > 0$ and $S, I, R, A \geq 0$, we know that eigenvalues satisfy these properties in the assumptions of our system for equilibrium. The first eigenvalue is 0 and $-\gamma$ and $-(\delta - \beta_2 kA) - \sqrt{D})/2$ are ≤ 0 as given by Eq. (21).

$$(-(\delta - \beta_2 kA) + \sqrt{D})/2 \leq 0 \tag{21}$$

Thus, rearranging and squaring both sides, we get to Eq. (22):

$$(\delta - \beta_2 kA)^2 \leq (\delta - \beta_2 kA)^2 + 4\beta_2\delta kA + 4\beta_2\delta kS \tag{22}$$

Hence, Eq. (22) holds true $\forall A, S \geq 0$, and at this point, the mathematical stability of the system is proved quantitatively by finding the eigenvector related to every eigenvalue. An equilibrium point is hyperbolic if none of the eigenvalues have zero real part. One of our eigenvalues is zero. Thus, the equilibrium is non-hyperbolic. If all eigenvalues have a negative real part, the equilibrium is steady, as given by the Lyapunov stability theorem (Lyapunov 1992). As evident in Eqs. (20–22), the corresponding eigenvalues are ≤ 0, and the equilibrium for our model is stable.

3.5 The simplification to information propagation SEI model

When the information is non-rumor, the model can directly be simplified to an information propagation model that leads to influence maximization achieved through anti-spreaders.

The roles of the nodes, however, remain almost the same. Anti-spreaders are still the carriers of correct information and the recovered nodes still are the dormant nodes that know about the information but are not spreading it. Similarly, susceptible nodes still remain the nodes that are unaware of the information and can likely be turned into recovered nodes by anti-spreaders. This effect can be seen when we plug in $I(t) = 0$, $\frac{d}{dt} I(t) = 0$ into the model (Eqs. 1, 2, 3 and 5) and we get the following system of equations, namely Eqs. (23 24, 25 and 26).

$$S(t) + A(t) + R(t) = 1 \tag{23}$$

$$\frac{d}{dt} A(t) = \delta R(t) \tag{24}$$

$$\frac{d}{dt} R(t) = -\delta R(t) + \beta_2 * k * S(t) * A(t) \tag{25}$$

$$\frac{d}{dt} S(t) = -\beta_2 * S(t) * A(t) * k \tag{26}$$

For the purpose of an analogy between these states, we rename the states for direct correspondence with the SEI model as follows:

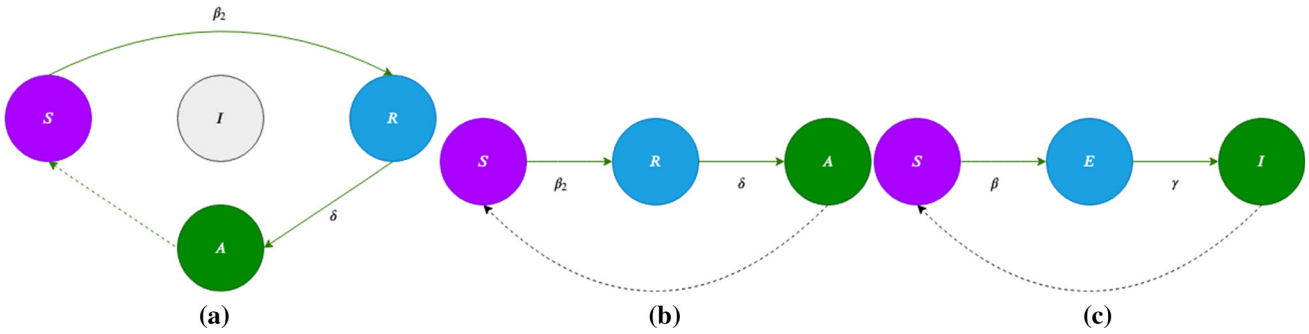


Fig. 2 Simplification of the model from $SIRA$ to SEI for non-rumor information

1. Anti-spreader nodes (A) become source spreaders or seed nodes (I)
2. Recovered nodes (R) are renamed as exposed nodes (E)
3. Susceptible nodes remain in the same naming convention (S)

These seed nodes (I) affect β_2 fraction of their neighbors, and they convert into exposed (E). Exposed users can become spreaders of the information with the rate δ.

Thus, the proposed $SIRA$ model for the rumor propagation and rumor control simplifies to information spreading SEI (Guihua and Zhen 2004) model for influence maximization. This conversion is depicted in Fig. 2

3.6 Simulation of the proposed model on a toy network

A toy network, as shown in Fig. 3, is chosen as a simulation network. The most influential node in the network is chosen as the spreader, and $\beta_1 = 0.5, \beta_2 = 1, \gamma = 1, \delta = 1$ are assumed without loss of generality. Initially, all the nodes, except the one chosen as the influential node based on out-degree, are susceptible, and the given node is in an Infected state. As time progresses, random neighbors of this node become infected, and the node recovers. After recovery, it is dormant, and other nodes continue spreading information. The recovered node decides to dispel the rumor and makes itself an anti-spreader. This effect makes all its susceptible neighbors instantly recover. The spreading continues till all the nodes either are against the rumor or have never heard of the rumor or its dispel. We describe the spread in Fig. 3 as a change of state in an event-driven simulation.

The change of state in an event-driven simulation are as follows:

t1: Node F is infected, chosen because of maximum out-degree, all other nodes are currently susceptible.

t2: Node F has recovered and has infected its neighbors with a uniform probability of $\beta_1 = 0.5$, thus randomly infecting nodes E and J.

t3: Node F is now an anti-spreader, all the infected nodes recover at the rate of $\gamma = 1$, and their neighbors are infected at the rate β_1.

t4: Node F has recovered its susceptible neighbor H at rate β_2, all recovered nodes became anti-spreaders at the rate of $\delta = 1$.

t5: Anti-spreaders are beginning to take more portion of the network. However, there is still an infection going on in the network.

t6: No neighbors of the current infected nodes are susceptible.

t7: All reachable nodes in the network have seen the information. However, there is still a node G, which has never heard of the rumor or anti-rumor.

t8: All nodes that saw the rumor or the truth are now anti-spreaders. Other nodes like G are still susceptible. After the equilibrium point, either all nodes will end as susceptible, or anti-spreaders and no nodes are spreading the rumor if the parameters $\beta_1, \beta_2, \gamma, \delta$ are coherent to the reproduction number.

In the above toy network simulation in Fig. 3, the value of various probabilities considered is $\beta_1 = 0.5, \beta_2 = 1, \gamma = 1, \delta = 1$, which are relatively very high values. The conversion rate from $S- > I$ is 0.5%, i.e., an infected node can randomly infect 50% of its susceptible neighbors. Similarly, each of the conversion rates from $I- > R$, $S- > R-$ and $R- > A$ is 1. It implies the probability of conversion from susceptible to recovered, infected to recovered and recovered to anti-spreaders is 100%. At the end of the simulation, all the nodes are in the anti-spreader state except one node in the susceptible state. The node 'G' is non-reachable from all other nodes, and hence, it is not influenced by any nodes during the simulation process and its color remains unchanged.

4 Result and analysis

In this section, we simulate our proposed methodology and analyze the different findings related to rumor propagation and rumor control using real-world various weighted and

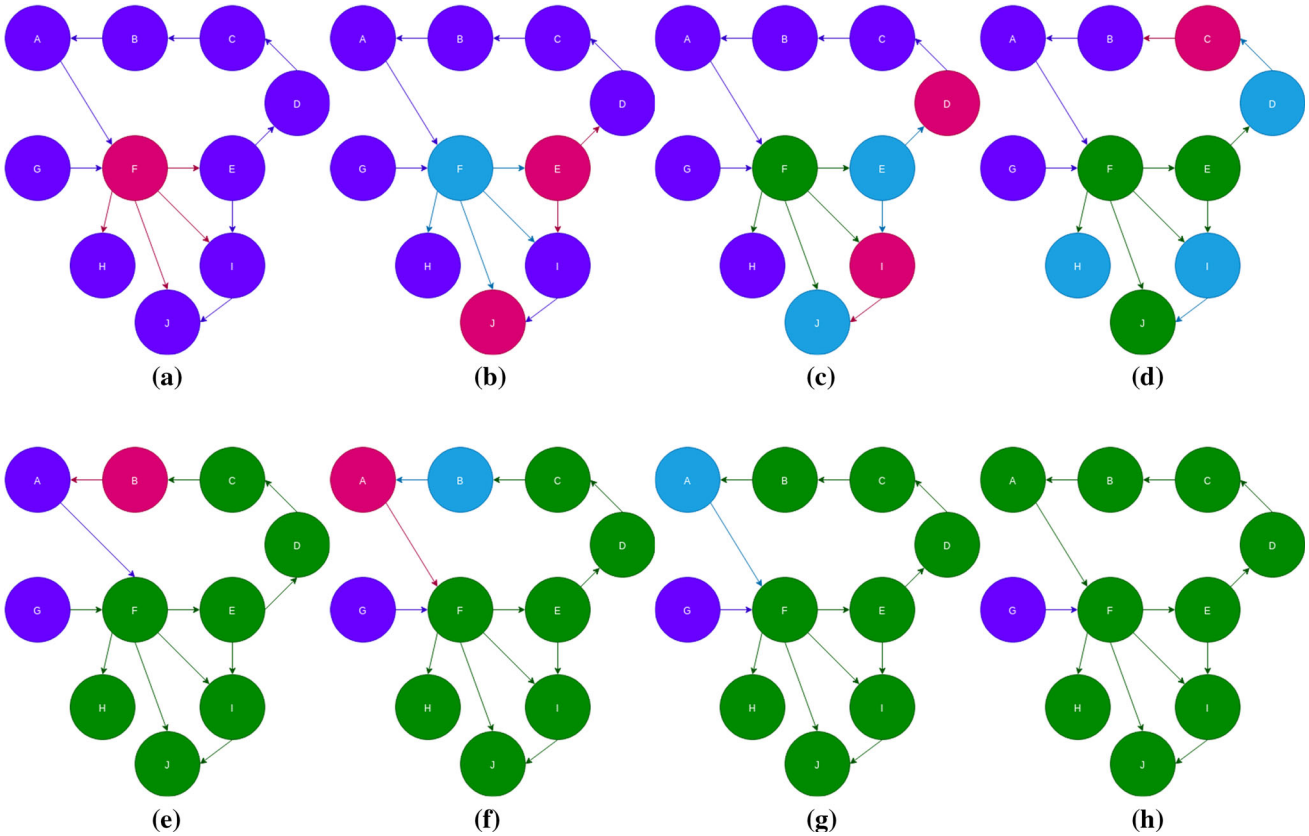


Fig. 3 All nodes are initially susceptible and an influential node is chosen as the spreader with $\beta_1 = 0.5$, $\beta_2 = 1$, $\gamma = 1$, $\delta = 1$. Here, purple nodes are susceptible, pink nodes are infected, blue nodes are recovered and green nodes are anti-spreaders

unweighted networks given in Subsection 4.1. Table 1 depicts the brief description used datasets in this paper. We use the spreading characteristics defined by $F(t)$ giving the count of the susceptible, infected, anti-spreaders and recovered nodes as a function of time. Infection scale, i.e., $F(t)$, is an important evaluation parameter. When information propagates in a network then at any time t, $F(t)$ denotes the fraction of nodes in the network which have received the particular information. A higher value of $F(t)$ indicates that information spreads rapidly in the network. Section 4.2 discusses the propagation of non-rumors, similar to the traditional SEI information propagation model, using the proposed $SIRA$ model. Section 4.3 discovers the role of anti-spreaders in containing the harmful effects of the rumor. In section 4.4, a novel method for comparison of centralities is proposed. Here, we compare the recently introduced NCVoteRank (Kumar and Panda 2020) centrality algorithm for unweighted networks and WVoteRank Kumar and Panda (2022) centrality on weighted networks for finding the initial spreaders. These centrality algorithms are simulated in comparison to traditional algorithms, DegreeRank Freeman (1978) for unweighted networks, and its weighted version WDegreeRank for weighted networks.

4.1 Datasets used

4.2 Information propagation on real-world networks

As highlighted in Section 3.5, the proposed algorithm reduces to a Susceptible-Exposed-Infected (SEI) model where susceptible nodes remain susceptible (S), recovered nodes act as exposed (E) nodes and anti-spreader nodes are the carrier of the correct information and act as infected (I) nodes. The information diffuses based on two parameters, namely the rate of infection as β_2 and rate of recovery as δ, leading to a simplified model of information propagation. As evident, the model has similar characteristics to those of the SEI model with similar parameters and has expected behavior on the real-world weighted and unweighted networks, as shown in Fig. 4. We use NCVoteRank centrality (Kumar and Panda 2020) and to identify one top spreader as infected node on the unweighted networks like Brightkite and Email-Enron. Similarly, we utilize WVoteRank centrality (Kumar and Panda 2022) to identify one top spreader as infected node to start the information diffusion process on the weighted networks like US-Airports and PowerGrid. Notice

Table 1 Description of datasets used

Dataset name	Characteristics	Description
US-Airports (Colizza et al. 2007)	Weighted, Directed, 1574 vertices, 28,236 edges	The network of the most crowded trading airports in the United States. An edge is there between airports, if there is a flight between them. The weights are the number of seats available on these flights
Brightkite (Cho et al. 2011)	Unweighted, Undirected, 58228 nodes, 214078 edges	Brightkite was a location-based social network. Users can share their locations by check-ins
Email-Enron (Leskovec et al. 2009)	Unweighted, Undirected, 36692 nodes, 183831 edges	Nodes of the network are email addresses. If address A sends one or more email to address B, the graph contains an undirected edge between these nodes
US PowerGrid (Watts and Strogatz 1998)	Undirected, Weighted, 4941 nodes, 6594 edges	This network is the high-voltage power grid in the USA. The nodes are substations, transformers and generators, and the connections are voltage power lines
CA-GrQc (Leskovec et al. 2007)	Unweighted, Undirected, 5242 nodes, 14496 edges	Arxiv GR-QC (General Relativity and Quantum Cosmology) research network is from the research papers and covers scientific referencing between papers in General Relativity and Quantum Cosmology category
p2p-Gnutella08 (Ripeanu et al. 2002)	Unweighted, Undirected, 6301 nodes, 20777 edges	This is a Gnutella peer-to-peer file-sharing network dataset collected from 08 August 2002 where a vertex represents host in the network, and the edge denotes the connection between the Gnutella hosts

that all the graphs behave similarly to the SEI epidemic model. As evident from Fig. 4a, the spreader on weighted network US-Airports is able to infect around 25% (750 nodes of the network receive the information at the end of the simulation, out of the total population of 3000) of the network. In Fig. 4b, the spreader on Brightkite is able to infect around 50% of the network and ended up having a stable equilibrium. In Fig. 4c, the spreader on Email-Enron is able to infect more than 50% of the network, notice that the peak reaches an early time stamp due to high connectivity. On the weighted PowerGrid network, notice that the change in the numbers is gradual, and very few nodes eventually become exposed as shown in Fig. 4d.

4.3 Numerical simulations for basic reproduction number and the equilibrium point

In this section, we discuss the equilibrium characteristics of the proposed method near the rumor-free equilibrium point, along with endemic equilibrium characteristics and the role of anti-spreaders nodes to minimize the effect of rumor on various network datasets. On reaching equilibrium on real-world networks, the rumor vanishes in the network only after the complete network has no more rumor-spreading nodes. The equilibrium is stable, only if the conditions are such that $R_t < 1$, as given by calculation in Eq. (10). We assume the values of the parameters β_1, β_2, and γ based on the reproduction number and network topology. The value of δ shall indicate how likely an informed (recovered) node is to dispel the rumor by becoming an anti-spreader. The value of δ depends on the preference of the given group and the severity of the information. As evident from Fig. 5a, the spreaders on weighted network US-Airports can infect around 35% of the network in the beginning, whereas around 50% nodes of the network recovers and becomes anti-spreader during the simulation and minimizing the spread of the rumor. In Fig. 5b, the spreader on Brightkite is able to infect around 15% of the network, and about 20% of the nodes became anti-spreaders. Due to the influence of anti-spreader, approximately 80% of the nodes remains susceptible means unaffected by the rumor. Hence, the roles of anti-spreaders nodes become vital to confine the rumor in the system. In Fig. 5c, the spreader on Email-Enron is able to infect about 32% of the network initially, and 43% nodes are anti-spreaders. It is clear that due to the impact of anti-spreaders about 70% of nodes remains susceptible. On the weighted PowerGrid network, the spreader is able to infect about 15% of the network initially, and 22% nodes are anti-spreaders, which enables approximately 70% nodes are susceptible and are unaffected by the rumor as depicted in Fig. 5d. Similarly, in the case of GrQc and p2p

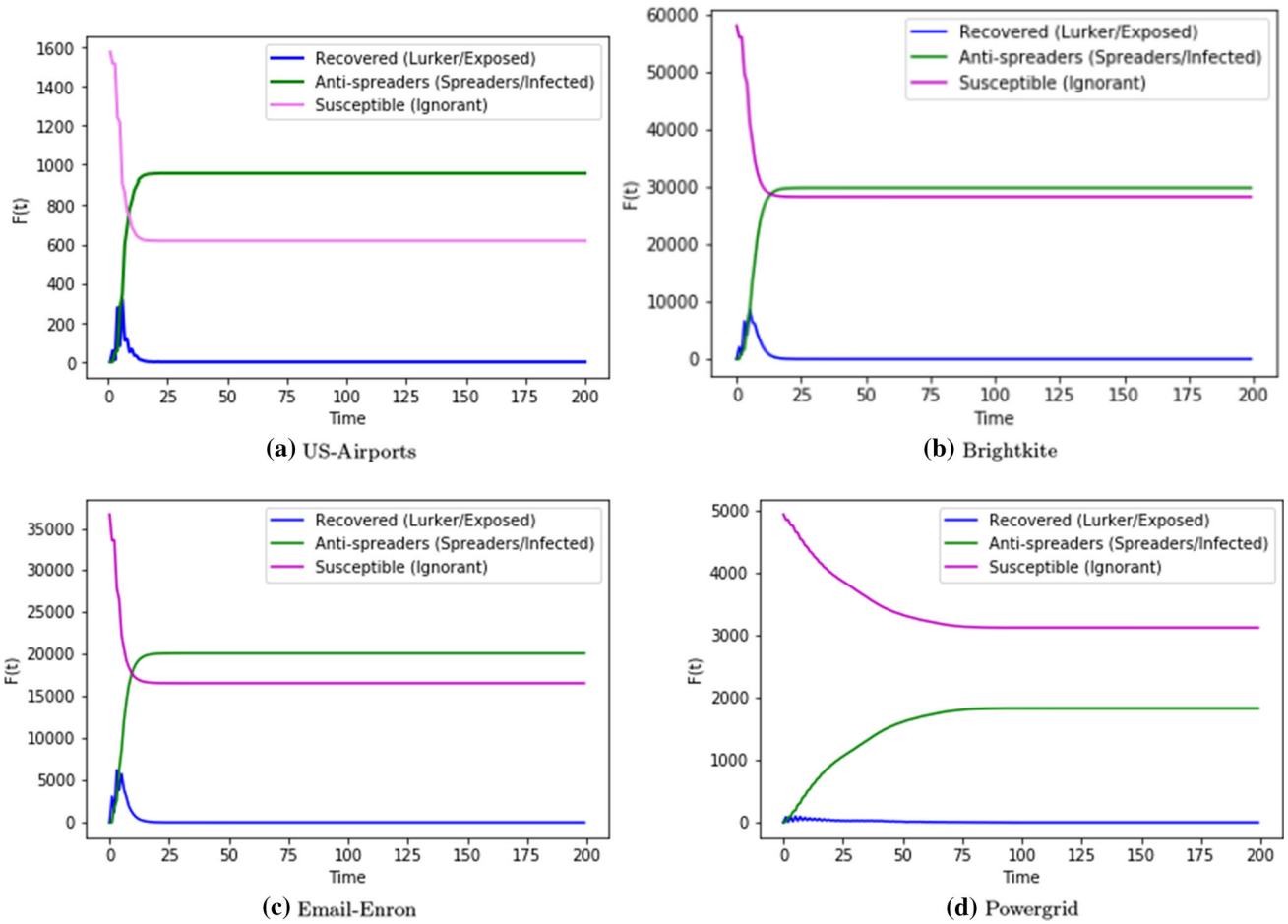


Fig. 4 Information propagation on real-world networks when there is no rumor, and anti-spreader nodes act as the information carrier. All nodes are initially susceptible. An influential node is chosen as the spreader using the NCVoteRank method in case of an unweighted network and WVoteRank method in weighted networks. The results are obtained over 100 independent simulation of SEI model with $\beta = 0.2$, and $\delta = 0.8$

Guntella08 networks, as evident from Fig. 5e and f, respectively, the anti-spreader nodes are able to control the rumor, and most of the nodes remain susceptible in the system after reaching the equilibrium.

Here, in case of the results obtained on real-life network datasets in Fig. 5, results are obtained over 100 independent simulation of SIRA model with the value of various probabilities as $\beta_1 = 0.1$, $\beta_2 = 0.1$, $\gamma = 0.2$, $\delta = 0.9$. The probability of conversion from $S- > I$ is 0.1, i.e., an infected node can randomly infect 10% of its susceptible neighbors. The probability of conversion $S- > I$ is 0.1, i.e., an infected node can randomly infect 10% of its susceptible neighbors. The probability of conversion from $S- > R$ is 0.1, which implies due to the influence of anti-spreader nodes, their 10% susceptible neighbors go directly in recovered state. Similarly, the rate of conversion from infected state to recovered state is 20%, and the rate of conversion from recovered state to anti-spreader state is 90%. Hence, in this case, we have performed the simulation on relatively lower conversion rates as compared to the simulation performed in the toy network as shown in Fig. 3. This results in a significant number of nodes that are in the susceptible state apart from in anti-spreader state when the equilibrium is reached in the system. The increase in the value of the conversion rate, especially β_1 for $S- > I$ and β_2 for $S- > R$, results in less number of susceptible nodes at the end of the simulation when equilibrium is reached. Also, many real-life networks like US-Airports, Brightkite and social networks are sparse networks where many nodes exist with relatively less degree. Here, degree means the number of direct neighbors or number of friends in social networks, and some of such nodes may not come in contact with either infected or anti-spreader nodes in the system and remains in the susceptible state or unaware of the information.

The endemic equilibrium state is the state where the disease or rumor cannot be totally eradicated but remains in

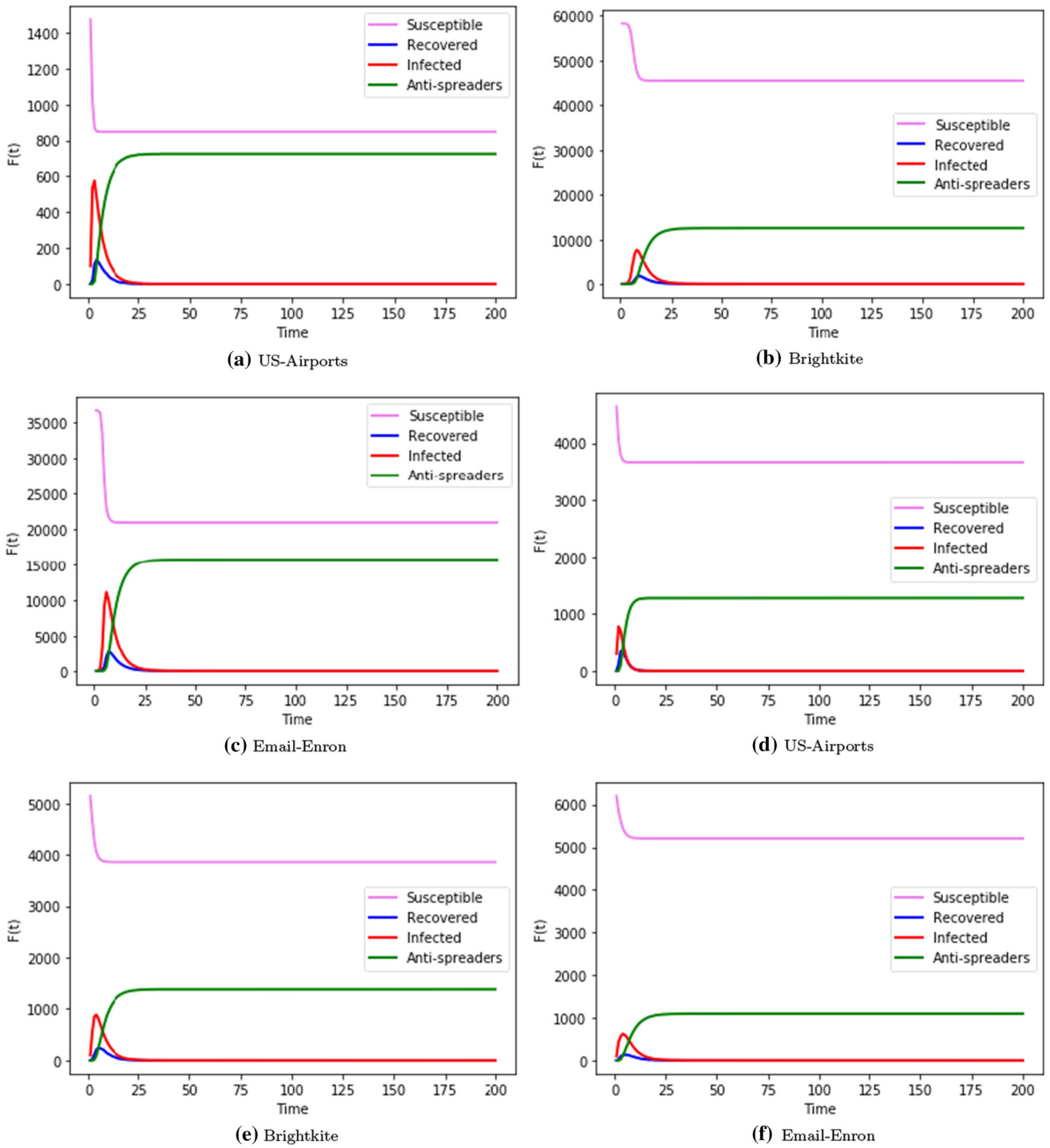


Fig. 5 Equilibrium characteristics of $SIRA$ model and impact of anti-spreader to control the rumor as seen on various used datasets. The results are obtained over 100 independent simulation of the proposed model with $\beta_1 = 0.1$, $\beta_2 = 0.1$, $\gamma = 0.2$, and $\delta = 0.9$

the population (Ugwa et al. 2013). For the disease or rumor to persist in the population, the anti-spreader class, the susceptible class, the infectious class and the recovered class must not be zero at the equilibrium state, i.e., $(S, I, R, A) \neq (0, 0, 0, 0)$. In this case the equilibrium is stable, only if the conditions are such that $R_t > 1$, as given by calculation in Eq. (10). The results obtained in Fig. 6 on the various used datasets confirm that there exit an endemic equilibrium point and its stability in the system.

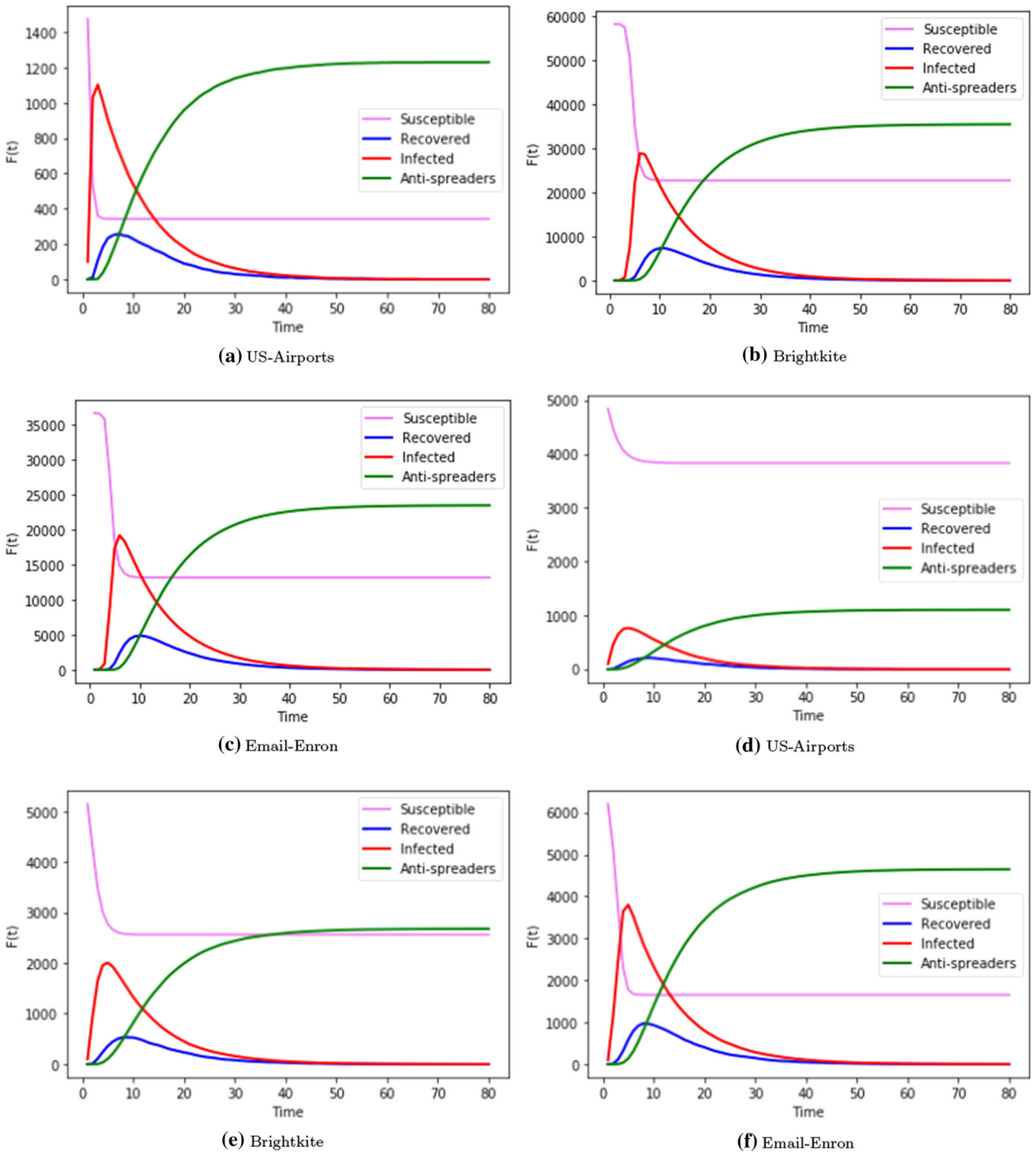
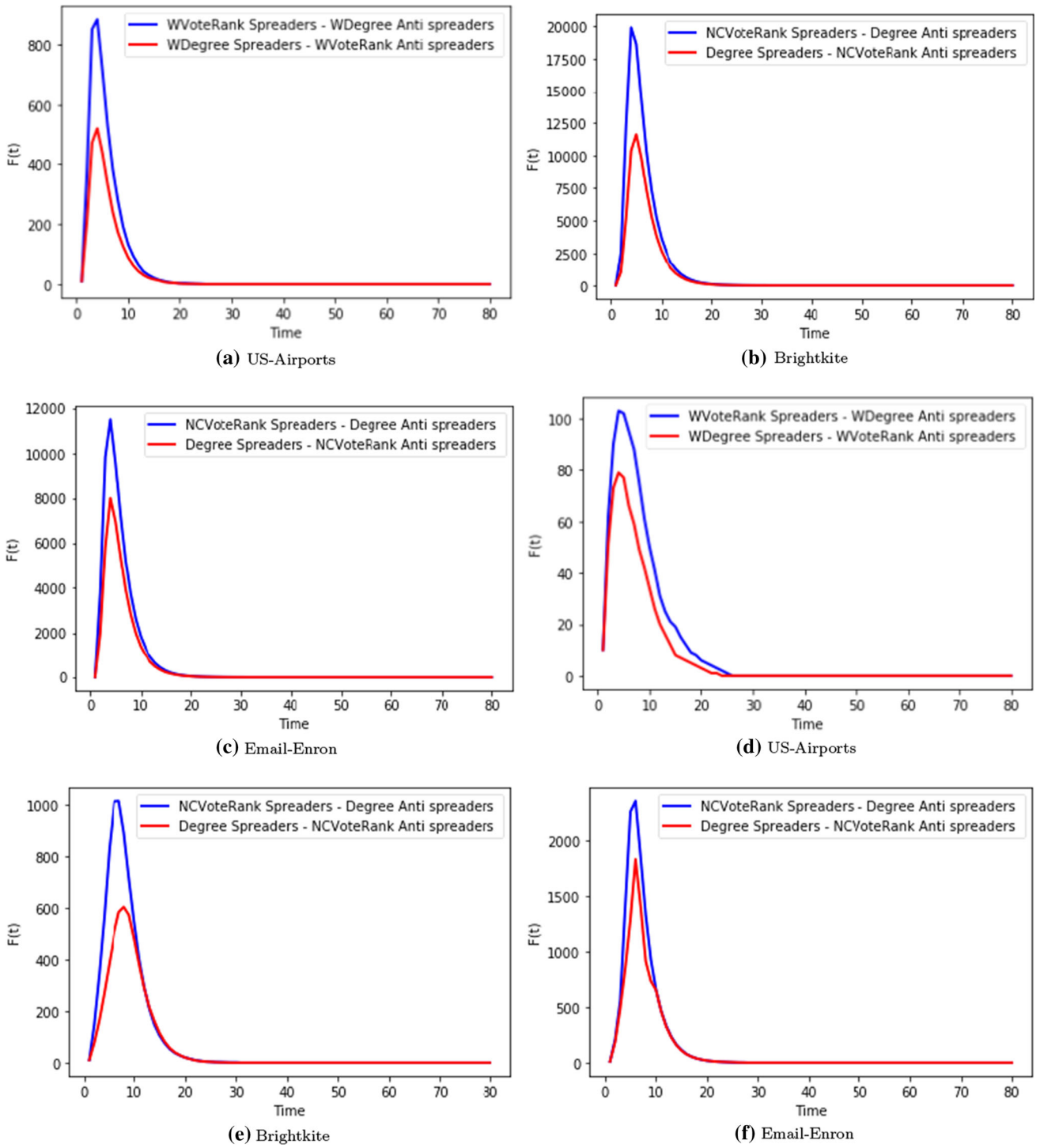


Fig. 6 Endemic equilibrium characteristics of $SIRA$ model observed on different used datasets. The results are obtained over 100 independent simulation of SIRA model with $\beta_1 = 0.4$, $\beta_2 = 0.01$, $\gamma = 0.1$, and $\delta = 0.3$



(a) US-Airports **(b)** Brightkite **(c)** Email-Enron **(d)** US-Airports **(e)** Brightkite **(f)** Email-Enron

Fig. 7 Spreading of the information using comparing centrality algorithms on the various network datasets with NCVoteRank and DegreeRank used on unweighted networks, and WVoteRank and WDegreeRank on weighted networks with parameters values as $\beta_1 = 0.4$, $\beta_2 = 0.3$, $\gamma = 0.3$, $\delta = 0.5$

4.4 Spreader and anti-spreader seeding: comparing centrality algorithms

The proposed $SIRA$ model can compare two different node centrality algorithms for evaluating the spread and control of information to model the real-life situation. In the framework, we can change the comparing centrality algorithms. Node centrality methods intend to assign some score to each node of the network based on its significance and different parameters (Kumar et al. 2022a, b) In this case, spreaders are chosen from one centrality algorithm spread the rumor, and the anti-spreaders are chosen from the second algorithm dispel the rumor and vice versa. We apply $F(t)$, the infected scale, as a function of time, as the performance measure for this comparison. We employ the top 10 selected spreaders and top 10 selected anti-spreaders using the specified centrality algorithm for performing the comparison study. We intend to develop a generic framework where any two node centrality can be taken to compare the performance. In the case of unweighted networks, we consider NCVoteRank and Degree centrality as the choice of node centrality algorithms. However, in the case of the weighted networks, WVoteRank and WDegree centrality is viewed as the choice of node centrality algorithms. During the simulation, due to a nonzero value of δ, the count of anti-spreaders increases, and also newly became anti-spreaders can immune their neighbors as per the dynamics of the model.

The obtained results of comparing centrality is depicted in Fig. 7 for various considered datasets by using the parameters $\beta_1 = 0.4, \beta_2 = 0.3, \gamma = 0.3, \delta = 0.5$. From the results, it can be inferred that initially, the count of infected nodes increases due to the influence of selected seed nodes. However, shortly the initial seed and the converted anti-spreader nodes tend to slow down the conversion of susceptible nodes to the infected node and rapidly decrease the count of spreaders by converting them to recovered nodes based on the parameters considered. PowerGrid network is a weighted network, and we use WVoteRank centrality and WDegreeRank centrality for the selection of spreaders and anti-spreaders, respectively, and vice versa. WVoteRank as the choice for spreaders and WDegreeRank as the choice for anti-spreaders can get 350% more infected users than WDegreeRank as the method for selecting spreaders and WVoteRank for the anti-spreaders, as evident in Fig. 7d. CA-GrQc is an unweighted network, and hence, NCVoteRank and DegreeRank are applicable for comparison. NCVoteRank as the choice for spreaders and DegreeRank as the choice for anti-spreaders are able to get 80% more infected users than DegreeRank as the method for the selection of spreaders and NCVoteRank for the anti-spreaders, as evident in Fig. 7e. Hence, we can conclude that NCVoteRank, on unweighted networks, and WVoteRank, on weighted networks, act as better node selection algorithms at both the spreading, and the control of the rumor and fake news as compared to degree on the unweighted and weighted degree on the weighted networks.

5 Conclusion

The healthcare social networks present potential risks for patients due to the possible distribution of poor quality or wrong information. In this paper, we proposed the $SIRA$ rumor propagation model to investigate the process of rumor propagation and control in healthcare social networks. We added a new state to the model called the anti-spreader that spread the truth, to prevent the harmful effects of the rumor. Considering the variety in people's reactions to rumors, we divided nodes into four categories, namely susceptible, anti-spreaders, recovered and infected. We modeled the system of differential equations mathematically and determined the reproduction number to find the equilibrium point and also confirmed the consistency of the proposed model on real-world datasets. We generalized our model into an information propagation model similar to the traditional models for modeling information diffusion to achieve influence maximization. In both weighted and unweighted networks, we compared two different centrality algorithms: one for choosing the spreaders and the other for choosing the anti-spreaders and vice versa and the amount of information spread. The studies showed that our model is coherent with the spread of rumors in social network topology, and our proposal gives a great reference to simulating and controlling the spread of rumors and fake news on the network. The proposed model is not applicable for heterogeneous networks. In the near future, we can extend our model to heterogeneous networks.

Author Contributions All the authors have contributed equally in the research, compiling results and manuscript preparation.

Funding No funding has been received.

Data availability Enquiries about data availability should be directed to the authors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The work conducted is not plagiarized. No one has been harmed in this work.

Consent to participate All the authors have given consent to submit the manuscript.

Consent for publication The authors provide their consent for the publication.

References

Britton T (2010) Stochastic epidemic models: a survey. Math Biosci 225(1):24–35

Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: friendship and mobility: user movement in location-based social networks. In: ACM SIGKDD international conference on knowledge discovery and data mining (KDD)

Colizza V, Pastor-Satorras R, Vespignani A (2007) Reaction-diffusion processes and metapopulation models in heterogeneous networks. Nat Phys 3:276–282

Freeman LC (1978) Centrality in social networks conceptual clarification. Soc Netw 1(3):215–239

Guihua L, Zhen J (2004) Global stability of an SEI epidemic model. Chaos, Solitons Fractals 21(4):925–931

Harko T, Lobo FSN, Mak MK (2014) Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and the SIR model with equal death and birth rates. Appl Math Comput 236(2014):184–194

Huang J, Jin X (2011) Preventing rumor spreading on small-world networks. J Syst Sci Complex 24(3):449–456

Indu V, Thampi SM (2019) The nature-inspired approach based on the forest fire model for modeling rumor propagation in social networks. J Netw Comput Appl 125:28–41

Izhikevich EM (2007) Equilibrium. Scholarpedia 2(10):2014

Kaur S, Kumar P, Kumaraguru P (2020) Automating fake news detection system using multi-level voting model. Soft Comput 24(12):9049–9069

Kijkuit B, Ende J (2007) The organizational life of an idea: integrating social network, creativity, and decision-making perspectives. J Manage Stud 44(6):863–882

Kumar S, Panda BS (2020) Identifying influential nodes in social networks: neighborhood coreness based voting approach. Stat Mech Appl, Phys A 553:124215

Kumar S, Panda A (2022) Identifying influential nodes in weighted complex networks using an improved WVoteRank approach. Appl Intell 52:1838–1852

Kumar S, Saini M, Goel M, Aggarwal N (2020) Modeling information diffusion in online social networks using SEI epidemic model. Procedia Comput Sci 171:672–678

Kumar S, Mallik A, Khetarpal A, Panda BS (2022) Influence maximization in social networks using graph embedding and graph neural network. Inf Sci. https://doi.org/10.1016/j.ins.2022.06.075

Kumar S, Gupta A, Khatri I (2022) CSR: A community based spreaders ranking algorithm for influence maximization in social networks. World Wide Web, 1–20

Lai K, Xiong X, Jiang X, Sun M, He L (2020) Who falls for rumor? Influence of personality traits on false rumor belief. Pers Individ Differ 152:109520

Leskovec J, Kleinberg J, Faloutsos C (2007) Graph evolution: densification and shrinking diameters. ACM Trans Knowl Discov Data (ACM TKDD) 1(1):2

Leskovec J, Lang K, Dasgupta A, Mahoney M (2009) Community structure in large networks: natural cluster sizes and the absence of large well-defined clusters. Internet Math 6(1):29–123

Lyapunov AM (1992) The general problem of the stability of motion. Int J Control 55(3):531–534

O'Keeffe GS, Clarke-Pearson K (2011) The impact of social media on children, adolescents, and families. Pediatrics 127(4):800–804

Pfeffer J, Zorbach T, Carley KM (2013) Understanding online firestorms: negative word-of-mouth dynamics in social media networks. J Mark Commun 20(1–2):117–128

Qiu L, Jia W, Niu W, Zhang M, Liu S (2021) SIR-IM: SIR rumor spreading model with influence mechanism in social networks. Soft Comput 25(22):13949–13958

Ripeanu M, Foster I, Iamnitchi A (2002) Mapping the gnutella network: properties of large-scale peer-to-peer systems and implications for system design. arXiv preprint cs/0209028

Roussel MR (2005) Nonlinear Dynamics, lecture notes. Stability analysis for ODEs. University Hall, Canada

Sivasankari S, Vadivu G (2021) Tracing the fake news propagation path using social network analysis. Soft Comput 1-9

Spohr D (2017) Fake news and ideological polarization. Bus Inf Rev 34(3):150–160

Tripathy RM, Bagchi A, Mehta S (2010) A study of rumor control strategies on social networks. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management - CIKM '10. Presented at the 19th ACM international conference, 1817–1820

Ugwa KA, Agwu IA, Akuagwu NA (2013) Mathematical analysis of the endemic equilibrium of the transmission dynamics of tuberculosis. Int J Technol Enhanc Emerg Eng Res 2(12):263–9

Wan YP, Zhang DG, Ren QH (2015) Propagation and inhibition of online rumor with considering the rumor elimination process. Acta Phys Sin 64(24):240501

Wang C, Wang G, Luo X, Li H (2019) Modeling rumor propagation and mitigation across multiple social networks. Phys A 535:122240

Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. Nature 393:440–442

Yang A, Huang X, Cai X, Zhu X, Lu L (2019) ILSR rumor spreading model with degree in a complex network. Phys A 531:121807

Zhu H, Ma J (2018) Analysis of SHIR rumor propagation in random heterogeneous networks with dynamic friendships. Phys A 513:257–271

Zhu L, Liu W, Zhang Z (2020) Delay differential equations modeling of rumor propagation in both homogeneous and heterogeneous networks with a forced silence function. Appl Math Comput 370:124925

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

arXiv:2208.01245v1 [math.CV] 2 Aug 2022

STARLIKE FUNCTIONS ASSOCIATED WITH A NON-MA-MINDA FUNCTION

S. SIVAPRASAD KUMAR AND POOJA YADAV

Abstract. In this work, we introduce a newly defined class of analytic functions associated with elliptical and strip domains, given by

$$\mathcal{F}[A,B] := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \psi_{A,B}(z) := \frac{1}{A-B} \log \frac{1+Az}{1+Bz} \right\},$$

where either $A = -B = \alpha$ or $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$ ($\alpha \in (0,1]$ and $\gamma \in (0, \pi/2]$). Here, we study various characteristic properties of $\psi_{A,B}(z)$ as well as functions in $\mathcal{F}[A,B]$. We also obtain the sharp radius of starlikeness of order δ and univalence for the functions in $\mathcal{F}[A,B]$. Further, we find the sharp radii for functions in $\mathcal{BS}(\alpha) := \{f \in \mathcal{A} : zf'(z)/f(z) - 1 \prec z/(1-\alpha z^2),\ \alpha \in (0,1)\}$, $\mathcal{S}_{cs}(\alpha) := \{f \in \mathcal{A} : zf'(z)/f(z) - 1 \prec z/((1-z)(1+\alpha z)),\ \alpha \in (0,1)\}$ and others to be in the class $\mathcal{F}[A,B]$.

1. Introduction

Let $\mathcal{A}$ denote the class of analytic functions f defined on the open unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, such that $f(z) = z + a_2 z^2 + a_3 z^3 + \cdots$. The subclass of $\mathcal{A}$ consisting of all univalent functions is denoted by $\mathcal{S}$. For two analytic functions f and g, we say that f is subordinate($\prec$) to g, if $f(z) = g(\omega(z))$ where $\omega(z)$ is a Schwarz function. Let $\mathcal{S}^*$ be the class of starlike functions, consists of functions $f \in \mathcal{S}$ such that $zf'(z)/f(z) \prec (1+z)/(1-z)$. In [6], Ma and Minda unifies all the subclasses of $\mathcal{S}^*$ as follows:

$$\mathcal{S}^*(\Phi) := \left\{ f \in \mathcal{S} : \frac{zf'(z)}{f(z)} \prec \Phi(z) \right\}, \tag{1.1}$$

where Φ is univalent with $\Phi'(0) > 0$ satisfying

A. $\Phi(z) \prec (1+z)/(1-z)$
B. $\Phi(\mathbb{D})$ is symmetric about real axis and starlike with respect to $\Phi(0) = 1$.

Further, Kumar and Banga [3], called such Φ as *Ma-Minda function* and the class of all Ma-Minda functions is denoted by $\mathcal{M}$. In [3], authors extensively studied Ma-Minda functions and classified them as *non-Ma-Minda function of type-A* whenever condition A doesn't hold, the class of all such functions is denoted by $\widetilde{\mathcal{M}}_A$. In past many authors studied the class (1.1) for various choices of $\Phi \in \mathcal{M}$, such as $\mathcal{S}^*(2/(1+e^{-z})) =: \mathcal{S}^*_{SG}$ [1], $\mathcal{S}^*(e^z) =: \mathcal{S}^*_e$ [9], $\mathcal{S}^*(z + \sqrt{1+z^2})$ [10], $\mathcal{S}^*(\sqrt{1+z}) =: \mathcal{S}^*_L$ [13], etc. Also the well known classes $\mathcal{S}^*(A,B)$ and $\mathcal{SS}^*(\alpha)$ are also obtained by taking $\Phi(z)(\in \mathcal{M})$ as $(1+Az)/(1+Bz)$ $(-1 \le B < A \le 1)$ and $((1+z)/(1-z))^\alpha$ $(\alpha \in (0,1])$ respectively. In particular, $\mathcal{S}^*(1,-1) =: \mathcal{S}^*$, $\mathcal{S}^*(1-2\delta,-1) =: \mathcal{S}^*(\delta)$ $(\delta \in [0,1))$, the class of starlike functions of order δ. Robertson [11] introduced and investigated the class of starlike functions of order δ, denoted by $\mathcal{S}^*_\delta$ for $\delta \le 1$. Note that if $\delta < 0$, then the functions in $\mathcal{S}^*_\delta$ may not be univalent, i.e. if $\delta < 0$, then $\Phi(z) = (1 + (1-2\delta)z)/(1-z) \in \widetilde{\mathcal{M}}_A$ and $\mathcal{S}^*_\delta \not\subset \mathcal{S}^*$. Later

2010 *Mathematics Subject Classification.* 30C45, 30C50, 30C80.

Key words and phrases. Analytic, Subordination, Logatithmic function, Starlike function, Non-univalent function.

many authors considered the classes associated with non-Ma-Minda functions of type-A such as the class by Uralegaddi [14]

$$\mathscr{M}(\beta)=\left\{f\in\mathcal{A}:\frac{zf'(z)}{f(z)}\prec\frac{1+(2\beta-1)z}{1-z},\ \beta>1\right\},$$

the classes associated with the Booth Lemniscate and Cissoid of Diocles

$$\mathcal{BS}^*(\alpha):=\left\{f\in\mathcal{A}:\left(\frac{zf'(z)}{f(z)}-1\right)\prec\frac{z}{1-\alpha z^2},\ \alpha\in[0,1)\right\},$$

$$\mathcal{S}^*_{cs}(\alpha):=\left\{f\in\mathcal{A}:\left(\frac{zf'(z)}{f(z)}-1\right)\prec\frac{z}{(1-z)(1+\alpha z)},\ \alpha\in[0,1)\right\},$$

which were introduced by [2] and [7] respectively. Similarly, authors in [8] studied the class $\mathcal{ST}^*_L(s):=\{f\in\mathcal{A}:zf'(z)/f(z)\prec\mathbb{L}_s(z)\}$ associated with the Limaçon of Pascal $\mathbb{L}_s(z)=(1+sz)^2$, $s\in[-1,1]\backslash\{0\}$. Note that $\mathbb{L}_s(z)\in\widehat{\mathcal{M}_A}$, whenever $|s|>1/\sqrt{2}$. Recently, Kumar and Gangania [5] studied the following class:

$$\mathcal{F}(\psi):=\left\{f\in\mathcal{A}:\frac{zf'(z)}{f(z)}-1\prec\psi(z)\right\},\tag{1.2}$$

where $\psi(z)\in\mathcal{S}^*$. If $1+\psi(z)\prec(1+z)/(1-z)$ and $1+\psi(\mathbb{D})$ is symmetric about the real axis with $\psi'(0)>0$ then $\mathcal{F}(\psi)$ reduces to the class $\mathcal{S}^*(1+\psi)$, a Ma-Minda class and if $1+\psi(z)\not\prec(1+z)/(1-z)$ then the functions in $\mathcal{F}(\psi)$ may not be univalent, thus $\mathcal{F}(\psi)\not\subseteq\mathcal{S}^*$. Note that all above mentioned classes associated with non-Ma-Minda functions of type-A are the special cases of $\mathcal{F}(\psi)$ such as $\mathcal{F}((1+(2\beta-1)z)/(1-z))=:\mathscr{M}(\beta)$, $\mathcal{F}(z/(1-\alpha z^2))=:\mathcal{BS}^*(\alpha)$, $\mathcal{F}(z/((1-z)(1+\alpha z)))=:\mathcal{S}^*_{cs}(\alpha)$ and $\mathcal{F}((1+sz)^2)=:\mathbb{L}_s(z)$.

Motivated by the above, we define a family of functions $\psi_{A,B}(z)$ as

$$\psi_{A,B}(z):=\frac{1}{A-B}\log\left(\frac{1+Az}{1+Bz}\right)\qquad z\in\mathbb{D},\tag{1.3}$$

where either $A=-B=\alpha$ or $A=\alpha e^{i\gamma}$ and $B=\alpha e^{-i\gamma}$, for $0<\alpha\leq1$ and $0<\gamma\leq\pi/2$. The function defined by (1.3) is analytic everywhere except the branchcuts $\{-\infty<\operatorname{Re}Az\leq-1,\ \operatorname{Im}Az=0\}\cup\{-\infty<\operatorname{Re}Bz\leq-1,\ \operatorname{Im}Bz=0\}$ with $\psi_{A,B}(0)=\psi'_{A,B}(0)-1=0$, i.e. $\psi_{A,B}(z)\in\mathcal{A}$. Now we introduce the following classes associated with $\psi_{A,B}(z)$:

Definition 1.1. Let $p\in\mathcal{S}$. Then $p\in\mathcal{L}[A,B]$ if and only if

$$p(z)\prec\psi_{A,B}(z).$$

Definition 1.2. Let $p\in\mathcal{A}$. Then $p\in\mathcal{F}[A,B]$ if and only if

$$\frac{zf'(z)}{f(z)}-1\prec\psi_{A,B}(z).\tag{1.4}$$

This paper aims to investigate various characteristic properties of functions in the classes $\mathcal{L}[A,B]$ and $\mathcal{F}[A,B]$. We also obtain the sharp radius of starlikeness of order δ and univalence for the functions in $\mathcal{F}[A,B]$. Further, we find the sharp radii for functions in $\mathcal{BS}(\alpha)$, $\mathcal{S}_{cs}(\alpha)$, etc to be in the class $\mathcal{F}[A,B]$.

2. Characteristics of $\mathcal{L}[A,B]$ and $\mathcal{F}[A,B]$

We now proceed to find various geometric properties of $\psi_{A,B}(z)$ and obtain certain bounds and inclusion relations for functions in the class $\mathcal{L}[A,B]$. Further we deduce the extremal function and

derive the growth and covering theorems for the class $\mathcal{F}[A,B]$. Note that

$$\psi_{A,B}(z)=\sum_{n=1}^{\infty}(-1)^{n+1}C_n z^n,$$

where

$$C_n=\frac{A^n-B^n}{n(A-B)}=\frac{1}{n}\sum_{k=0}^{n-1}A^kB^{n-1-k}\qquad(n=1,2,\ldots). \tag{2.1}$$

Remark 2.1. From (2.1) we have

(i) Since C_n is real valued for all n, $\psi_{A,B}(\mathbb{D})$ is symmetric with respect to real axis.

(ii) $|C_n|=\left|\frac{1}{n}\sum_{k=0}^{n-1}A^kB^{n-1-k}\right|\leq\alpha^{n-1}\leq 1.$

Theorem 2.2. *The function $\psi_{A,B}(z)$ is convex and univalent on $\mathbb{D}$.*

Proof. Let $H(z):=1+z\psi''_{A,B}(z)/\psi'_{A,B}(z)$. By [12, Corollary 3], it is enough to show that $\operatorname{Re}H(z)>0$. Since $|A|=|B|=\alpha\leq 1$, using basic calculation with [4, Theorem 1], we obtain

$$\begin{aligned}\operatorname{Re}H(z)&=\operatorname{Re}\left(-1+\frac{1}{1+Az}+\frac{1}{1+Bz}\right)\\&>-1+\frac{2}{1-\alpha}>0.\end{aligned}$$

Thus the result holds. ∎

Theorem 2.3. *Let $p\in\mathcal{L}[A,B]$, Then in the disk $\mathbb{D}_r=\{z\in\mathbb{C}:|z|\leq r<1\}$,*

(i) For $A=-B=\alpha$, we have

$$|\operatorname{Re}p(z)|\leq\frac{1}{2\alpha}\log\left(\frac{1+\alpha r}{1-\alpha r}\right) \tag{2.2}$$

$$|\operatorname{Im}p(z)|\leq\frac{1}{2\alpha}\sin^{-1}\left(\frac{2\alpha r}{1+\alpha^2r^2}\right). \tag{2.3}$$

(ii) For $A=\alpha e^{i\gamma}$ and $B=\alpha e^{-i\gamma}$, we have

$$\frac{-1}{2\alpha\sin\gamma}(\eta-\tau)\leq\operatorname{Re}p(z)\leq\frac{1}{2\alpha\sin\gamma}(\eta+\tau). \tag{2.4}$$

$$\frac{1}{2\alpha\sin\gamma}\log T_2\leq\operatorname{Im}p(z)\leq\frac{1}{2\alpha\sin\gamma}\log T_1. \tag{2.5}$$

where $\eta:=\sin^{-1}\left(\frac{2\alpha r\sin\gamma}{\sqrt{1+\alpha^4r^4-2\alpha^2r^2\cos2\gamma}}\right)$, $\tau:=\tan^{-1}\left(\frac{-\alpha^2r^2\sin2\gamma}{1-\alpha^2r^2\cos2\gamma}\right)$ *and* $T_j:=\left(\frac{\sqrt{1+\alpha^4r^4-2\alpha^2r^2\cos2\gamma}+(-1)^j2\alpha r\sin\gamma}{1-\alpha^2r^2}\right)^{-1}$ $(j=1,2)$.

Proof. Let w be the Schwarz function such that $w(0)=0$ and $|w(z)|\leq|z|=r<1$ for all $z\in\mathbb{D}$. Define the function $F_w:\mathbb{D}\to\mathbb{C}$ such that $p(z)=\psi_{A,B}(w(z))=\frac{1}{A-B}\log\left(F_w(z)\right)$. We have

$$F_w(z):=\frac{1+Aw(z)}{1+Bw(z)},$$

which maps unit disk into the disk

$$F_w(\mathbb{D}) \subset \left\{\zeta \in \mathbb{C}; \left|\zeta - \frac{1-A\overline{B}}{1-|B|^2}\right| < \frac{|A-B|}{1-|B|^2}\right\}.$$

As $|A|= \alpha \in (0,1]$, therefore origin $O \notin F_w(\mathbb{D})$ (for more details see [4]). Now using [4, Theorem 1] with $|z|= r < 1$, we have

$$\frac{|1-A\overline{B}r^2|-|A-B|r}{1-|B|^2r^2} \leq |F_w(z)| \leq \frac{|1-A\overline{B}r^2|+|A-B|r}{1-|B|^2r^2}, \tag{2.6}$$

$$\left|\arg F_w(z) - \tan^{-1}\frac{\mathrm{Im}(A\overline{B})r^2}{\mathrm{Re}(A\overline{B})r^2-1}\right| < \sin^{-1}\frac{|A-B|r}{|1-A\overline{B}r^2|}. \tag{2.7}$$

Since $\log f(z) = \log|f(z)|+i\arg f(z)$ and $p(z) = \dfrac{1}{A-B}\log\left(F_w(z)\right)$, thus for $A = -B = \alpha$ we obtain (2.2) and (2.3) and for $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$, we have (2.4) and (2.5). ∎

For the sake of computational convenience, we shall make the following assumptions:

$$h_1 := \frac{1}{2\alpha}\log\left(\frac{1+\alpha}{1-\alpha}\right), \qquad h_2 := \frac{1}{2\alpha}\sin^{-1}\left(\frac{2\alpha}{1+\alpha^2}\right),$$

$$k_1 := \frac{1}{2\alpha\sin\gamma}\sin^{-1}\left(\frac{2\alpha\sin\gamma}{\sqrt{1+\alpha^4-2\alpha^2\cos 2\gamma}}\right), \qquad k_2 := \frac{1}{2\alpha\sin\gamma}\log\frac{\sqrt{1+\alpha^4-2\alpha^2\cos 2\gamma}+2\alpha\sin\gamma}{1-\alpha^2}$$

and

$$k := \frac{1}{2\alpha\sin\gamma}\tan^{-1}\left(\frac{\alpha^2\sin 2\gamma}{\alpha^2\cos 2\gamma - 1}\right).$$

Now we shall discuss the image domain of $\psi_{A,B}$ for different cases of α:
For $0 < \alpha < 1$, $\psi_{A,B}(\mathbb{D}) = \{u+iv \in \mathbb{C} : (u,v) \in \Omega_1\}$ where

$$\Omega_1 := \begin{cases} \dfrac{u^2}{h_1^2} + \dfrac{v^2}{h_2^2} < 1, \text{when } A = -B = \alpha \\ \dfrac{(u-k)^2}{k_1^2} + \dfrac{v^2}{k_2^2} < 1, \text{when } A = \alpha e^{i\gamma}, B = \alpha e^{-i\gamma} \text{ for } 0 < \gamma \leq \pi/2 \end{cases}. \tag{2.8}$$

Moreover when $\alpha = 1$, the major axis h_1 and k_2 respectively, tends to infinity, hence $\psi_{A,B}(\mathbb{D}) = \{u+iv \in \mathbb{C} : (u,v) \in \Omega_2\}$ where Ω_2 is the following strip domain

$$\Omega_2 := \begin{cases} -\dfrac{\pi}{4} < v < \dfrac{\pi}{4}, & \text{when } A = -B = 1 \\ \dfrac{\gamma-\pi}{2\sin\gamma} < u < \dfrac{\gamma}{2\sin\gamma}, & \text{when } A = e^{i\gamma}, B = e^{-i\gamma} \text{ for } 0 < \gamma \leq \pi/2 \end{cases}. \tag{2.9}$$

Remark 2.4. 1. $\max_{|z|=r}\mathrm{Re}\,\psi_{A,B}(z) = \psi_{A,B}(r)$ and $\min_{|z|=r}\mathrm{Re}\,\psi_{A,B}(z) = \psi_{A,B}(-r)$.
2. If $\alpha < 1$, then $p \prec \psi_{A,B}(z)$ if and only if $p(\mathbb{D}) \subseteq \Omega_1$ and if $\alpha = 1$, then $p \prec \psi_{A,B}(z)$ if and only if $p(\mathbb{D}) \subseteq \Omega_2$, where Ω_1 and Ω_2 are as defined in (2.8) and (2.9) respectively.
3. $1 + \psi_{A,B} \in \widetilde{\mathcal{M}}_A$.

Let $w = 1 + \psi_{A,B}(z)$, then we have

$$|w-1| = |\psi_{A,B}(z)| = \frac{1}{|A-B|}\left|\log\left(\frac{1+Az}{1+Bz}\right)\right|.$$

It can be easily seen that for $\alpha \in (0,1)$, $A = -B = \alpha$

$$\min_{|z|=1}|\psi_{A,B}(z)| = h_2 \text{ and } \max_{|z|=1}|\psi_{A,B}(z)| = h_1$$

and for $A=\alpha e^{i\gamma}$ and $B=\alpha e^{-i\gamma}$

$$\min_{|z|=1}|\psi_{A,B}(z)|= k_1+k \text{ and } \max_{|z|=1}|\psi_{A,B}(z)|= k_2.$$

Remark 2.5. For $\alpha=1$, h_1 and k_2 tends to ∞, $h_2=\pi/4$ and $k_1+k=\gamma/(2\sin\gamma)$.

Thus we have the following inclusion relation:

Lemma 2.6. *Let $\alpha\in(0,1)$ and $\gamma\in(0,\pi/2]$, then we have*

(i) for $A=-B=\alpha$,

$$\{w\in\mathbb{C}:|w-1|<h_2\}\subset 1+\psi_{A,B}(\mathbb{D})\subset\{w\in\mathbb{C}:|w-1|<h_1\}.$$

(ii) for $A=\alpha e^{i\gamma}$ and $B=\alpha e^{-i\gamma}$,

$$\{w\in\mathbb{C}:|w-1|<k_1+k\}\subset 1+\psi_{A,B}(\mathbb{D})\subset\{w\in\mathbb{C}:|w-1|<k_2\}.$$

Further for $\alpha=1$, we have

(iii) for $A=-B=1$,

$$\left\{w\in\mathbb{C}:|w-1|<\frac{\pi}{4}\right\}\subset 1+\psi_{A,B}(\mathbb{D}).$$

(iv) for $A=e^{i\gamma}$ and $B=e^{-i\gamma}$,

$$\left\{w\in\mathbb{C}:|w-1|<\frac{\gamma}{2\sin\gamma}\right\}\subset 1+\psi_{A,B}(\mathbb{D}).$$

In general, if we consider a disc $\mathcal{D}(a,r):=\{w\in\mathbb{C};|w-a|<r\}$ where $a\in\mathbb{R}$ and $1\in\mathcal{D}(a,r)$, then we observe that $\mathcal{D}(a,r)\subset 1+\psi_{A,B}(\mathbb{D})$ if and only if

$$\mathcal{D}(a,r)\subset\begin{cases}\mathcal{D}(h_2,1), & \text{for } A=-B=\alpha,\\ \mathcal{D}(k_1+k,1), & \text{for } A=\alpha e^{i\gamma},\ B=\alpha e^{-i\gamma},\end{cases}\tag{2.10}$$

i.e. if and only if

$$|a-1|\leq\begin{cases}h_2-r, & \text{for } A=-B=\alpha,\\ k_1+k-r, & \text{for } A=\alpha e^{i\gamma},\ B=\alpha e^{-i\gamma}.\end{cases}\tag{2.11}$$

Thus we have the following theorem:

Theorem 2.7. *Let $-1<D<C\leq 1$ and $p(z):=(1+Cz)/(1+Dz)$. Then $p(z)\prec\mathcal{L}[A,B]$ if and only if*

(i) for $A=-B=\alpha$

$$C\leq\begin{cases}h_2+(1-h_2)D, & \textit{when } (1-CD)/(1-D^2)\leq 1,\\ h_2+(1+h_2)D, & \textit{when } (1-CD)/(1-D^2)\geq 1.\end{cases}\tag{2.12}$$

(ii) for $A=\alpha e^{i\gamma}$ and $B=\alpha e^{-i\gamma}$,

$$C\leq\begin{cases}k_1+k+(1-k_1-k)D, & \textit{when } (1-CD)/(1-D^2)\leq 1,\\ k_1+k+(1+k_1+k)D, & \textit{when } (1-CD)/(1-D^2)\geq 1.\end{cases}\tag{2.13}$$

Proof. Since $p(z)=(1+Cz)/(1+Dz)$ maps $\mathbb{D}$ onto $\mathcal{D}(a,r)$, where

$$a=\frac{1-CD}{1-D^2}\quad\text{and}\quad r=\frac{C-D}{1-D^2},$$

the result follows at once from (2.10) and (2.11). ∎

From Theorem 2.7, we deduce the following corollary, which ensures that $\mathcal{F}[A,B]$ is nonempty.

Corollary 2.8. *Let $-1 < D < C \leq 1$ and $f \in \mathcal{A}$ be such that $zf'(z)/f(z) = (1+Cz)/(1+Dz)$, then $f \in \mathcal{F}[A,B]$ if and only if, the condition (2.12) or (2.13) holds.*

In particular, we illustrate below that the class $\mathcal{F}[A,B]$ is non-empty.

Example 1. 1. *If $f(z) = z\exp(Cz)$, then $f(z) \in \mathcal{F}[A,B]$ if and only if*

$$C \leq \begin{cases} h_2, & \text{for } A = -B = \alpha, \\ k_1 + k, & \text{for } A = \alpha e^{i\gamma} \ \& \ B = \alpha e^{-i\gamma}. \end{cases}$$

2. *If $f(z) = z/(1+Cz)$, then $f(z) \in \mathcal{F}[A,B]$ if and only if*

$$C \geq \begin{cases} -h_2/(1+h_2), & \text{for } A = -B = \alpha, \\ -(k_1+k)/(1+k_1+k), & \text{for } A = \alpha e^{i\gamma} \ \& \ B = \alpha e^{-i\gamma}. \end{cases}$$

3. *If $f(z) = z/(1-Cz)^2$, then $f(z) \in \mathcal{F}[A,B]$ if and only if*

$$C \leq \begin{cases} h_2/(2+h_2), & \text{for } A = -B = \alpha, \\ (k_1+k)/(2+k_1+k), & \text{for } A = \alpha e^{i\gamma} \ \& \ B = \alpha e^{-i\gamma}. \end{cases}$$

Now from (1.4), we have $f \in \mathcal{F}[A,B]$ if and only if there exists $p(z) \prec \psi_{A,B}(z)$ such that

$$f(z) = z\exp\left(\int_0^z \frac{p(t)}{t}dt\right). \tag{2.14}$$

If we take $p(z) = \psi_{A,B}(z)$, then we obtain from (2.14)

$$f_{A,B}(z) := z\exp\frac{Li_2(-Bz) - Li_2(-Az)}{A-B}, \tag{2.15}$$

where $Li_2(x) = \sum_{n=1}^{\infty} x^n/n^2$, denotes the Spence's (or dilogarithm) function. The function $f_{A,B}(z)$ is nonunivalent but have the extremal properties for many problems for the class $\mathcal{F}[A,B]$ (see Fig:1). Kumar and Gangania [5] obtained various geometric properties such as the growth and covering theorem for the case when $1+\psi(z) \not\prec (1+z)/(1-z)$. As a consequence of the same and Remark 2.4, we obtain the following result:

Theorem 2.9. *Let $f \in \mathcal{F}[A,B]$ and $f_{A,B}$ be defined as in (2.15), then for $|z| = r$*
(i) Growth Theorem: $-f_{A,B}(-r) \leq |f(z)| \leq f_{A,B}(r)$.
(ii) Covering Theorem: Either f is a rotation of $f_{A,B}$ or $\{w \in \mathbb{C} : |w| \leq -f_{A,B}(-1)\} \subset f(\mathbb{D})$.

For $z = re^{i\theta}$, where θ is fixed but arbitrary, as a consequence of growth theorem and $\psi_{A,B}(-r) \leq \operatorname{Re}\psi_{A,B}(re^{i\theta}) \leq \psi_{A,B}(r)$, we obtain

$$\log\frac{f(z)}{z} = \int_0^r \frac{p(te^{i\theta})}{t}dt,$$

where $p(z) := \psi_{A,B}(w(z))$ and w is a Schwarz functon. Further

$$\frac{f(z)}{z} = \exp\int_0^r \frac{p(te^{i\theta})}{t}dt = \exp\left(\int_0^r \operatorname{Re}\frac{p(te^{i\theta})}{t}dt + i\int_0^r \operatorname{Im}\frac{p(te^{i\theta})}{t}dt\right)$$

$$\left|\frac{f(z)}{z}\right| \leq \exp\int_0^r \frac{\operatorname{Re}\psi_{A,B}(te^{i\theta})}{t}dt$$

$$\exp\int_0^r \frac{\psi_{A,B}(-t)}{t}dt \leq \left|\frac{f(z)}{z}\right| \leq \exp\int_0^r \frac{\psi_{A,B}(t)}{t}dt.$$

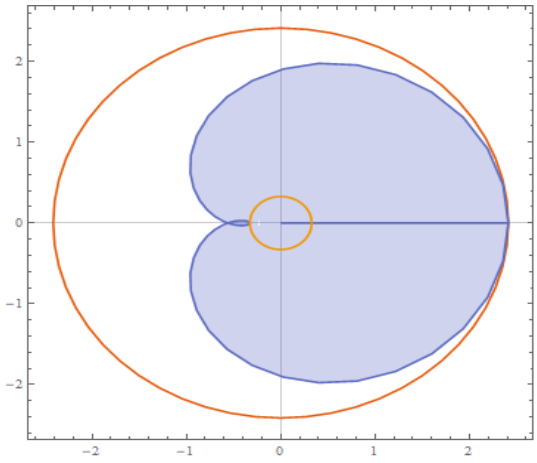


FIGURE 1. The images of $f_{A,B}(1)\partial\mathbb{D}$, $f_{A,B}(\mathbb{D})$ and $f_{A,B}(-1)\partial\mathbb{D}$, for $A = 0.5e^{i\pi/3}$ and $B = 0.5e^{-i\pi/3}$.

Corollary 2.10. *Let $f \in \mathcal{F}[A,B]$ and $M(r) = \exp\int_0^r \frac{\psi_{A,B}(t)}{t}dt$, then for $|z| = r$, we have*
(i) $M(-r) \leq \left|\frac{f(z)}{z}\right| \leq M(r)$.
(ii) $(1+\max_{|z|\leq r}|\psi_{A,B}(z)|)M(-r) \leq |f'(z)| \leq (1+\max_{|z|\leq r}|\psi_{A,B}(z)|)M(r)$.
(iii) $2\pi r(1+\max_{|z|\leq r}|\psi_{A,B}(z)|)M(-r) \leq L(f,r) \leq 2\pi r(1+\max_{|z|\leq r}|\psi_{A,B}(z)|)M(r)$.
(iv) $f(z)/z \prec f_{A,B}(z)/z$.

3. Radius Estimates

In this section, we obtain the sharp radius of univalence and radius of starlikeness of order δ for functions in $\mathcal{F}[A,B]$. Further, we find the sharp radii for functions in $\mathcal{BS}(\alpha)$, $\mathcal{S}_{cs}(\alpha)$ and others to be in the class $\mathcal{F}[A,B]$. From Theorem 2.3 we conclude that $f \in \mathcal{S}[A,B]$ is starlike of order $1+\frac{1}{|A-B|}(\varphi-\vartheta) < 0$, hence f may not be univalent in $\mathbb{D}$, thus $\mathcal{F}[A,B] \not\subseteq \mathcal{S}^*$. Therefore, in the following result, we find the radius of starlikeness of order δ, where $\delta \in [0,1)$ for functions in the class $\mathcal{F}[A,B]$.

Theorem 3.1. *Let $\alpha \in (0,1]$, $\gamma \in (0,\gamma_0)$, where $\gamma_0 \simeq 1.2461\ldots$ and $\delta \in [0,1)$ be given numbers. If $f \in \mathcal{F}[A,B]$, then f is starlike of order δ in the disc $|z| < r(\delta)$, where*

(i) For $A = -B = \alpha$,

$$r(\delta) = \frac{\exp(2\alpha(1-\delta))-1}{\alpha(\exp(2\alpha(1-\delta))+1)}. \tag{3.1}$$

(ii) For $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$, $r(\delta)$ is the smallest positive root of

$$\tan(2\alpha\sin\gamma(\delta-1)) - \frac{\alpha^4r^4\sin 2\gamma + 2\alpha^3r^3\sin\gamma\cos 2\gamma - \alpha^2r^2\sin 2\gamma - 2\alpha r\sin\gamma}{\alpha^4r^4\cos 2\gamma - 2\alpha^3r^3\sin\gamma\sin 2\gamma - \alpha^2r^2\cos 2\gamma - \alpha^2r^2 + 1} = 0. \tag{3.2}$$

The result is sharp.

Proof. Let $f \in \mathcal{F}[A,B]$, then using Theorem 2.3 for $|z| < r$, we have
Case 1. when $A = -B = \alpha$,

$$\operatorname{Re}\frac{zf'(z)}{f(z)} > 1 - \frac{1}{2\alpha}\log\left(\frac{1+\alpha r}{1-\alpha r}\right) =: h_1(r),$$

Now as $h_1'(r) < 0$ for all choices of α, thus $h_1(r)$ is strictly decreasing function from 1 to $1 - 1/(2\alpha)\log((1+\alpha)/(1-\alpha)) < 0$. Hence the root $r(\delta)$, given in (3.1), of the equation $h_1(r) = \delta$, is the radius of starlikeness of order δ of $\mathcal{F}[A, B]$.

Case 2. when $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$

$$\operatorname{Re}\frac{zf'(z)}{f(z)} > 1 - \frac{1}{2\alpha\sin\gamma}(\eta - \tau) =: h_2(r),$$

where η and τ are given in Theorem 2.3. Since $h_2(r) = \delta$ can be simplified to the equation (3.2) by a simple computation, thus it is enough to find the smallest positive root of (3.2). Further

$$h_2'(r) = -\frac{1 - \alpha^4 r^4 + 2\alpha r\cos\gamma(1-\alpha^2 r^2)}{(1-\alpha^2 r^2)\sqrt{1+\alpha^4 r^4 - 2\alpha^2 r^2\cos 2\gamma}} < 0$$

for all choices of α and γ, $h_2(r)$ is a strictly decreasing function from 1 to $g(\alpha, \gamma)$, where

$$g(\alpha,\gamma) := 1 + \frac{1}{2\alpha\sin\gamma}\left(\tan^{-1}\left(\frac{\alpha^2\sin 2\gamma}{\alpha^2\cos 2\gamma - 1}\right) - \sin^{-1}\left(\frac{2\alpha\sin\gamma}{\sqrt{1+\alpha^4 - 2\alpha^2\cos 2\gamma}}\right)\right).$$

Note that $g(\alpha, \gamma)$ is real valued, non-constant analytic function on a bounded domain $R := (0, 1) \times$

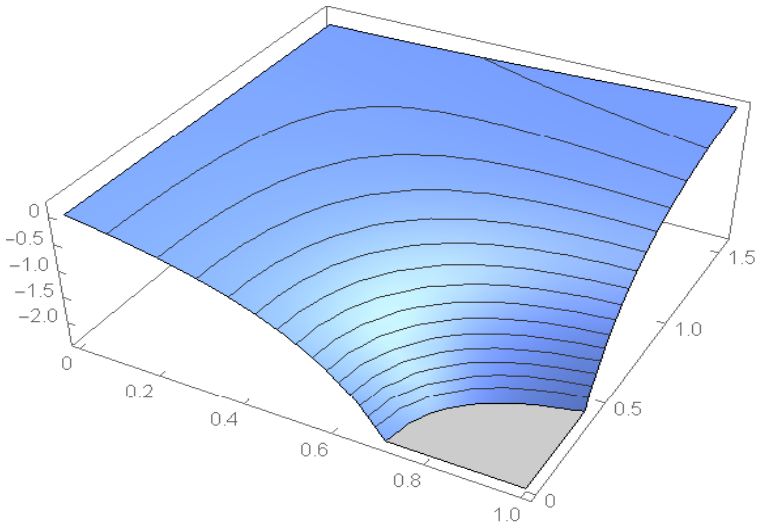


FIGURE 2. Figure of $g(\alpha, \gamma)$

$(0, \pi/2)$ and is bounded above on R. Thus by Maximum Modulus Theorem, $g(\alpha, \gamma)$ attains its maximum value on the boundary of R and the four boundaries are enlisted below:

1. $g(0, \gamma) = 0$.
2. $g(1, \gamma) = 1 + (\gamma - \pi)/(2\sin\gamma)$. Now as $g'(1, \gamma) = (1 + (\pi - \gamma)\cot\gamma)/(2\sin\gamma) > 0$, therefore $g(1, \gamma)$ is strictly increasing from $-\infty$ to $1 - \pi/4 \simeq 0.2146\ldots$ and γ_0 is the zero of it.
3. $g(\alpha, 0) = 1 - 1/(1-\alpha)$. As $g'(\alpha, 0) = \log(1-\alpha) < 0$, therefore $g(\alpha, 0)$ is strictly decreasing from 0 to $-\infty$.
4. $g(\alpha, \pi/2) = 1 - \left(\sin^{-1}\left(2\alpha/(1+\alpha^2)\right)\right)/(2\alpha)$. As $g'(\alpha, \pi/2) = \left(\left(2\alpha/(1+\alpha^2)\right) + \sin^{-1}\left(2\alpha/(1+\alpha^2)\right)\right)/(2\alpha^2) > 0$, therefore $g(\alpha, \pi/2)$ is also strictly increasing from 0 to $1 - \pi/4 \simeq 0.2146$.

From Figure 2 and all four boundaries, we observe that $g(\alpha, \gamma) < 0$ for all choices of $\alpha \in (0, 1]$, $\gamma \in (0, \gamma_0)$, where γ_0 is the root of the equation $g(1, \gamma) = 0$. Hence $r(\delta)$ is the radius of starlikeness of order δ of $\mathcal{F}[A, B]$. ∎

Theorem 3.2. *Let $\alpha \in (0,1]$ and $\delta \in [0,1)$. If $f \in \mathcal{F}[A,B]$, for the case when $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$, then there is an unique $\gamma' \in (\gamma_0, \pi/2)$ such that*

(i) when $\gamma \in (\gamma_0, \gamma']$, f is starlike of order δ in $\mathbb{D}_{r(\delta)}$, where $r(\delta)$ the root of the equation (3.2).
(ii) when $\gamma \in (\gamma', \pi/2)$, f is starlike of order $\delta \in [0, g(\alpha, \pi/2))$ in $\mathbb{D}$ and starlike of order $\delta \in [g(\alpha,\pi/2), 1)$ in $\mathbb{D}_{r(\delta)}$.

Proof. By observing the Figure 2 and the nature of all boundary curves discussed in Theorem 3.1, we arrive at $g(\alpha, \gamma_0) < 0$ and $g(\alpha, \pi/2) > 0$ for all $\alpha \in (0,1)$. Thus by IVP for any choice of $\alpha \in (0,1)$, there is an unique $\gamma' \in (\gamma_0, \pi/2)$ such that $g(\alpha, \gamma') = 0$, $g(\alpha, \gamma) < 0$ for all $\gamma \in (\gamma_0, \gamma')$ and $g(\alpha, \gamma) > 0$ for all $\gamma \in (\gamma', \pi/2)$. Hence the result follows. ∎

Taking $\delta = 0$ in Theorems 3.1 and 3.2, we obtain the following result:

Corollary 3.3. *Let $\alpha \in (0,1]$ and $\gamma \in (0, \pi/2]$. If $f \in \mathcal{F}[A,B]$ then*

(i) For $A = -B = \alpha$, f is starlike univalent in the disc $|z| < r_0$, where

$$r_0 = \frac{\exp(2\alpha) - 1}{\alpha(\exp(2\alpha) + 1)}.$$

(ii) For $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$, $f \in \mathcal{S}^*$, whenever $\gamma \in (\gamma', \pi/2)$, and f is starlike univalent in the disc $|z| < r_0$, whenever $\gamma \in (0, \gamma_0) \cup (\gamma_0, \gamma']$, where γ_0, γ' are given in Theorems 3.1 and 3.2, respectively and r_0 is the smallest positive root of

$$\tan(2\alpha \sin\gamma) + \frac{\alpha^4 r^4 \sin 2\gamma + 2\alpha^3 r^3 \sin\gamma \cos 2\gamma - \alpha^2 r^2 \sin 2\gamma - 2\alpha r \sin\gamma}{\alpha^4 r^4 \cos 2\gamma - 2\alpha^3 r^3 \sin\gamma \sin 2\gamma - \alpha^2 r^2 \cos 2\gamma - \alpha^2 r^2 + 1} = 0.$$

The result is sharp.

Corollary 3.4. *Let $f \in \mathcal{F}[A,B]$. Then for the disc $|z| < r \le r_0$, where r_0 is given in Corollary 3.3, we have*

(i) For $A = -B = \alpha$,

$$\left| \arg\left(\frac{zf'(z)}{f(z)} \right) \right| < \tan^{-1} \left(\frac{\sin^{-1}\left(\frac{2\alpha r}{1+\alpha^2 r^2}\right)}{2\alpha - \log\left(\frac{1+\alpha r}{1-\alpha r}\right)} \right)$$

(ii) For $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$,

$$\left| \arg\left(\frac{zf'(z)}{f(z)} \right) \right| < \tan^{-1} \left(\frac{\log T_2}{2\alpha \sin\gamma - \eta + \tau} \right),$$

where T_2, η and τ are given in Theorem 2.3.

Corollary 3.5. *Let $f \in \mathcal{F}[A,B]$. Then for the disc $|z| < r_s \le r_0$, where r_0 is given in Corollary 3.3, $f \in \mathcal{SS}^*(\beta)$, where $r_s \in (0, r_0]$ is the smallest positive root of the following equation:*

(i) For $A = -B = \alpha$,

$$\sin^{-1}\left(\frac{2\alpha r}{1+\alpha^2 r^2}\right) = \tan\left(\frac{\beta\pi}{2}\right)\left(2\alpha - \log\left(\frac{1+\alpha r}{1-\alpha r}\right)\right)$$

(ii) For $A = \alpha e^{i\gamma}$ and $B = \alpha e^{-i\gamma}$,

$$\frac{\log T_2}{2\alpha \sin\gamma - \eta + \tau} = \tan\left(\frac{\beta\pi}{2}\right),$$

where T_2, η and τ are given in Theorem 2.3.

We now discuss the radius estimates for the classes $\mathcal{BS}^*(\alpha)$ and $\mathcal{S}^*_{cs}$, which are not contained in $\mathcal{S}^*$. Further we derive radii for functions in $\mathcal{BS}^*(\alpha)$ and $\mathcal{S}^*_{cs}$ to be in $\mathcal{F}[A,B]$.

Theorem 3.6. *Let $\alpha \in (0,1)$, $r_0 = (-1+\sqrt{1+4\alpha h_1^2})/(2\alpha h_1)$ and $\alpha_0 \in (0,1)$ is the smallest solution of*

$$1 - \frac{r_0^2\cos^2\theta(1-\alpha r_0^2)^2}{(1+\alpha^2 r_0^4 - 2\alpha r_0^2\cos 2\theta)^2 h_1^2} - \frac{r_0^2\sin^2\theta(1+\alpha r_0^2)^2}{(1+\alpha^2 r_0^4 - 2\alpha r_0^2\cos 2\theta)^2 h_2^2} = 0,$$

where $\theta \in (0,\pi/2)$. If $f \in \mathcal{BS}^*(\alpha)$ then for the disc $|z| \leq r_b < 1$, $f \in \mathcal{F}[A,B]$, where

$$r_b = \begin{cases} r_0, & \text{for } A = -B = \alpha \leq \alpha_0 \\ r_1 := \dfrac{-1+\sqrt{1+4\alpha h_2^2}}{2\alpha h_2}, & \text{for } A = -B = \alpha > \alpha_0 \\ r_2 := \dfrac{-1+\sqrt{1+4\alpha(k_1+k)^2}}{2\alpha(k_1+k)}, & \text{for } A = \alpha e^{i\gamma} \text{ and } B = \alpha e^{-i\gamma}. \end{cases}$$

Moreover the radius r_0 and r_2 are sharp.

Proof. Let $f \in \mathcal{BS}^*(\alpha)$, then we have

Case 1. When $A = -B = \alpha$: Since $\max_{|z|=1} \operatorname{Re}\psi_{A,B}(z) = h_1$, to find such $r < 1$ for which the image of $zf'(z)/f(z) - 1$ under the disc $|z| < r$ lies inside $\psi_{A,B}(\mathbb{D})$, it is necessary that

$$\max_{|z|=r<1} \operatorname{Re}\left(\frac{z}{1-\alpha z^2}\right) = \frac{r}{1-\alpha r^2} \leq h_1 \tag{3.3}$$

should hold. Clearly for $|z| \leq r_0$, the inequality (3.3) holds. Now to see that for $\alpha \leq \alpha_0$, the radius r_0 is also sufficient for $zf'(z)/f(z) - 1 \prec z/(1-\alpha z^2) \in \psi_{A,B}(\mathbb{D})$ in the disc $|z| \leq r_0$. For $\zeta = re^{i\theta}$ ($\theta \in [0, 2\pi)$), we have

$$B_r(\theta) := \frac{\zeta}{1-\alpha\zeta^2} = \frac{r\cos\theta(1-\alpha r^2)}{1+\alpha^2 r^4 - 2\alpha r^2\cos 2\theta} + i\frac{r\sin\theta(1+\alpha r^2)}{1+\alpha^2 r^4 - 2\alpha r^2\cos 2\theta}.$$

Since $\operatorname{Re} B_r(\theta) = \operatorname{Re} B_r(-\theta)$, $\operatorname{Re} B_r(\theta) = -\operatorname{Re} B_r(\pi-\theta)$ and $\operatorname{Im} B_r(\theta) = \operatorname{Im} B_r(\pi-\theta)$, therefore the curve $B_r(\theta)$ is symmetric about real and imaginary axis thus it is sufficient to consider for $\theta \in [0, \pi/2]$. Now for $r = r_0$, the square of the distance from the origin to the points of $B_{r_0}(\theta)$ is given by

$$Dist(0; B_{r_0}(\theta)) := \frac{r_0^2}{1+\alpha^2 r_0^4 - 2\alpha r_0^2\cos 2\theta}.$$

Since $Dist(0; B_{r_0}(\theta))' < 0$, thus $Dist(0; B_{r_0}(\theta))$ is a decreasing function of θ. Hence the farthest point of $B_{r_0}(\theta)$ from origin is $(r_0/(1-\alpha r_0^2), 0)$, which lies on the boundary of Ω_1. Now for $A = -B = \alpha > \alpha_0$ and $A = \alpha e^{i\gamma}$, $B = \alpha e^{-i\gamma}$ with $|z| = r$, we have

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{r}{1-\alpha r^2}.$$

Therefore, by Lemma 2.6, we see that $\mathcal{F}[A,B]-$radius for the class $\mathcal{BS}^*(\alpha)$ are the smallest positive roots r_1 and r_2 of the equations $r/(1-\alpha r^2) = h_2$ and $r/(1-\alpha r^2) = k_1 + k$, respectively. ∎

Theorem 3.7. *Let $\alpha \in (0,1)$, $r_0 = \dfrac{-(1+(1-\alpha)h_1) + \sqrt{(1+(1-\alpha)h_1)^2 + 4\alpha h_1^2}}{2\alpha h_1}$ and $\alpha_0 \in (0,1)$ is the smallest solution of*

$$1 - \frac{r_0^2((\alpha-1)r_0 + \cos\theta(1-\alpha r_0^2))^2}{(1+r_0^2 - 2r_0\cos\theta)^2(1+\alpha^2 r_0^2 + 2\alpha r_0\cos\theta)^2 h_1^2} - \frac{r_0^2(1+\alpha r_0^2)^2\sin^2\theta}{(1+r_0^2 - 2r_0\cos\theta)^2(1+\alpha^2 r_0^2 + 2\alpha r_0\cos\theta)^2 h_2^2} = 0,$$

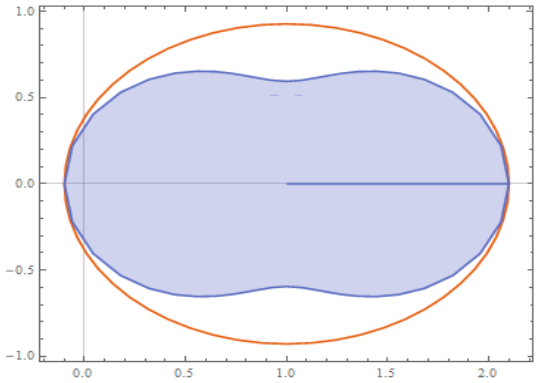


FIGURE 3. The image of $\partial\mathbb{D}$ and $\mathbb{D}_{r_0=0.772\ldots}$ under $\psi_{0.5,-0.5}(z)$ and $z/(1-0.5z^2)$, respectively.

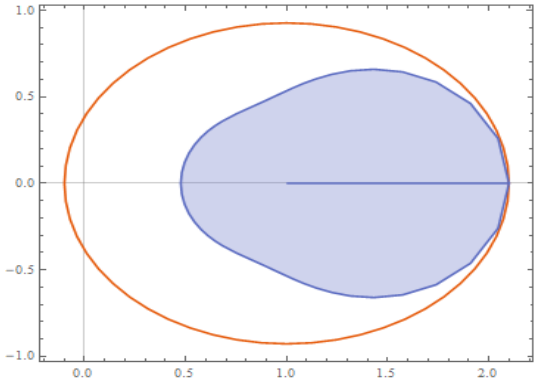


FIGURE 4. The image of $\partial\mathbb{D}$ and $\mathbb{D}_{r_0=0.587\ldots}$ under $\psi_{0.5,-0.5}(z)$ and $z/((1-z)1+0.5z)$, respectively.

where $\theta \in (0, \pi/2)$. *If* $f \in \mathcal{S}^*_{cs}(\alpha)$ *then for the disc* $|z| \leq r_{cs} < 1$, $f \in \mathcal{F}[A, B]$, *where*

$$r_{cs} = \begin{cases} r_0, & \text{for } A = -B = \alpha \leq \alpha_0 \\ r_1 := \dfrac{-(1+(1-\alpha)h_2) + \sqrt{(1+(1-\alpha)h_2)^2 + 4\alpha h_2^2}}{2\alpha h_2}, & \text{for } A = -B = \alpha > \alpha_0 \\ r_2 := \dfrac{-(1+(1-\alpha)(k_1+k)) + \sqrt{(1+(1-\alpha)(k_1+k))^2 + 4\alpha(k_1+k)^2}}{2\alpha(k_1+k)}, & \text{for } A = \alpha e^{i\gamma}, B = \alpha e^{-i\gamma}. \end{cases}$$

Moreover the radius r_0 *and* r_2 *are sharp.*

Proof. Let $f \in \mathcal{S}^*_{cs}(\alpha)$. Since for $A = -B = \alpha$, $\max_{|z|=1} \operatorname{Re} \psi_{A,B}(z) = h_1$. Thus to find such $r < 1$ for which the image of $zf'(z)/f(z) - 1$ under the disc $|z| < r$ lies inside $\psi_{A,B}(\mathbb{D})$, it is necessary that

$$\max_{|z|=r<1} \operatorname{Re}\left(\frac{z}{(1-z)(1+\alpha z)}\right) = \frac{r}{(1-r)(1+\alpha r)} \leq h_1 \tag{3.4}$$

must hold. Clearly for $|z| \leq r_0$, the equation (3.4) holds. Now to see that for $\alpha \leq \alpha_0$ radius r_0 is also sufficient for $zf'(z)/f(z) - 1 \prec z/((1-z)(1+\alpha z)) \in \psi_{A,B}(\mathbb{D})$ in the disc $|z| \leq r_0$. For $\zeta = re^{i\theta}$ $(\theta \in [0, 2\pi))$, we have

$$CS_r(\theta) := \frac{\zeta}{(1-\zeta)(1+\alpha\zeta)}$$
$$= \frac{r((\alpha-1)r+\cos\theta(1-\alpha r_0^2))}{(1+r^2-2r\cos\theta)(1+\alpha^2r^2+2\alpha r\cos\theta)} + i\frac{r(1+\alpha r_0^2)\sin\theta}{(1+r^2-2r\cos\theta)(1+\alpha^2r^2+2\alpha r\cos\theta)}.$$

Since $\operatorname{Re} CS_r(\theta) = \operatorname{Re} CS_r(-\theta)$, therefore the curve $CS_r(\theta)$ is symmetric about real axis thus it is sufficient to consider for $\theta \in [0,\pi]$. Now for $r = r_0$, the square of the distance from the origin to the points of $CS_{r_0}(\theta)$ is given by

$$Dist(0; CS_{r_0}(\theta)) := \frac{r_0^2}{(1+r^2-2r\cos\theta)^2(1+\alpha^2r^2+2\alpha r\cos\theta)^2}.$$

Since $Dist(0; CS_{r_0}(\theta))' = 0$ for $\theta = 0, \theta_0$ and π with $Dist(0; CS_{r_0}(\theta))' < 0$ whenever $\theta \in (0,\theta_0)$ and $Dist(0; CS_{r_0}(\theta))' > 0$ whenever $\theta \in (\theta_0, \pi)$. And $Dist(0; CS_{r_0}(0)) - Dist(0; CS_{r_0}(\pi)) > 0$, hence the farthest point of $CS_{r_0}(\theta)$ from origin is equal to h_1 obtained $\theta = 0$. Now for $A = -B = \alpha > \alpha_0$ and $A = \alpha e^{i\gamma}$, $B = \alpha e^{-i\gamma}$ with $|z| = r$, we have

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{r}{(1-r)(1+\alpha r)}.$$

Therefore, by Lemma 2.6, we see that $\mathcal{F}[A,B]-$radius for the class $\mathcal{S}^*_{cs}(\alpha)$ are the smallest positive roots r_1 and r_2 of the equations $r/((1-r)(1+\alpha r)) = h_2$ and $r/((1-r)(1+\alpha r)) = k_1 + k$, respectively. ∎

References

[1] P. Goel and S. Sivaprasad Kumar, Certain class of starlike functions associated with modified sigmoid function, Bull. Malays. Math. Sci. Soc. **43** (2020), no. 1, 957–991.

[2] R. Kargar, A. Ebadian and J. Sokół, On Booth lemniscate and starlike functions, Anal. Math. Phys. **9** (2019), no. 1, 143–154.

[3] S. S. Kumar and S. Banga, On a Special Type of Ma-Minda Function. arxiv preprint arXiv:2006.02111, (2020).

[4] S. S. Kumar and P. Yadav, On Oblique Domains of Janowski Function. arXiv preprint arXiv:2104.08533 (2021).

[5] S. S. Kumar and G. Kamaljeet, On Geometrical Properties of Certain Analytic functions, Iran. J. Sci. Technol. Trans. A Sci. (2020), https://doi.org/10.1007/s40995-021-01116-1.

[6] W. C. Ma and D. Minda, A unified treatment of some special classes of univalent functions, in *Proceedings of the Conference on Complex Analysis (Tianjin, 1992)*, 157–169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge.

[7] V. S. Masih, A. Ebadian and S. Yalçin, Some properties associated to a certain class of starlike functions, Math. Slovaca **69** (2019), no. 6, 1329–1340.

[8] V. S. Masih and S. Kanas, Subclasses of Starlike and Convex Functions Associated with the Limaçon Domain, Symmetry **12** (2020), no. 6, 942.

[9] R. Mendiratta, S. Nagpal and V. Ravichandran, On a subclass of strongly starlike functions associated with exponential function, Bull. Malays. Math. Sci. Soc. **38** (2015), no. 1, 365–386.

[10] R. K. Raina and J. Sokół, Some properties related to a certain class of starlike functions, C. R. Math. Acad. Sci. Paris **353** (2015), no. 11, 973–978.

[11] M. S. Robertson, Certain classes of starlike functions, Michigan Math. J. **32** (1985), no. 2, 135–140.

[12] K. Sakaguchi, A note on p-valent functions, J. Math. Soc. Japan **14** (1962), 312–321.

[13] J. Sokół and J. Stankiewicz, Radius of convexity of some subclasses of strongly starlike functions, Zeszyty Nauk. Politech. Rzeszowskiej Mat. No. 19 (1996), 101–105.

[14] B. A. Uralegaddi, M. D. Ganigi and S. M. Sarangi, Univalent functions with positive coefficients, Tamkang J. Math. **25** (1994), no. 3, 225–230.

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: spkumar@dce.ac.in

Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India
Email address: poojayv100@gmail.com

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/362711337

Starlike Functions Associated with a Petal Shaped Domain

Article *in* Bulletin of the Korean Mathematical Society · July 2022

DOI: 10.4134/BKMS.b210602

CITATION
1

READS
16

2 authors:

Kush Arora
Delhi Technological University
2 PUBLICATIONS **1** CITATION
SEE PROFILE

Sivaprasad Kumar
Delhi Technological University
74 PUBLICATIONS **665** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Coefficient Estimates and Differential Subordinations for certain class of analytic functions View project

All content following this page was uploaded by Kush Arora on 16 August 2022.

The user has requested enhancement of the downloaded file.

Bull. Korean Math. Soc. **59** (2022), No. 4, pp. 993–1010
https://doi.org/10.4134/BKMS.b210602
pISSN: 1015-8634 / eISSN: 2234-3016

STARLIKE FUNCTIONS ASSOCIATED WITH A PETAL SHAPED DOMAIN

Kush Arora and S. Sivaprasad Kumar

Abstract. In this paper, we establish some radius results and inclusion relations for starlike functions associated with a petal-shaped domain.

1. Introduction

Let the open unit disk $\{z \in \mathbb{C} : |z| < 1\}$ be represented by $\mathbb{D}$ and $\mathcal{H}$ be the class of all analytic functions in $\mathbb{D}$. Consider $\mathcal{A}_n$ as the class of analytic functions f in $\mathbb{D}$ represented by

$$f(z) = z + a_{n+1}z^{n+1} + a_{n+2}z^{n+2} + \cdots . \tag{1}$$

In particular, denote $\mathcal{A}_1 := \mathcal{A}$ and let $\mathcal{S}$ be the subclass of $\mathcal{A}$ such that it involves all univalent functions $f(z)$ in $\mathbb{D}$. Let g, h be two analytic functions and ω be a Schwarz function satisfying $\omega(0) = 0$ and $|\omega(z)| \leq |z|$ such that $g(z) = h(\omega(z))$ then g is said to be subordinate to h, or $g \prec h$. If h is univalent, then $g \prec h$ if and only if $g(0) = h(0)$ and $g(\mathbb{D}) \subset h(\mathbb{D})$. Ma and Minda [11] introduced the univalent function ψ satisfying $\operatorname{Re}\psi(\mathbb{D}) > 0$, $\psi(\mathbb{D})$ starlike with respect to $\psi(0) = 1$ and $\psi'(0) > 0$ and the domain $\psi(\mathbb{D})$ being symmetric about the real axis. Further, they gave the definitions for the general subclasses of starlike and convex functions, respectively, as follows:

$$\mathcal{S}^*(\psi) := \{f \in \mathcal{S} : zf'(z)/f(z) \prec \psi(z)\} \tag{2}$$

and

$$\mathcal{K}(\psi) := \{f \in \mathcal{S} : 1 + zf''(z)/f'(z) \prec \psi(z)\}. \tag{3}$$

For different choices of ψ, many subclasses of $\mathcal{S}^*$ and $\mathcal{K}$ can be obtained. For example, the notable classes of Janowski starlike and convex functions [8] are represented by $\mathcal{S}^*[C, D] := \mathcal{S}^*((1 + Cz)/(1 + Dz))$ and $\mathcal{K}[C, D] := \mathcal{K}((1 + Cz)/(1+Dz))$ for $-1 \leq D < C \leq 1$, respectively. Further, $\mathcal{S}^*_\alpha := \mathcal{S}^*[1-2\alpha, -1]$ and $\mathcal{K}_\alpha := \mathcal{K}^*[1-2\alpha, -1]$ represent the classes of starlike and convex functions

Received August 15, 2021; Accepted January 24, 2022.
2020 *Mathematics Subject Classification.* Primary 30C80; Secondary 30C45.
Key words and phrases. Starlike function, convex function, petal shaped domain, radius problems.

©2022 Korean Mathematical Society

of order $\alpha \in [0,1)$, respectively. Note that $\mathcal{S}^* := \mathcal{S}_0^*$ and $\mathcal{K} := \mathcal{K}_0$ represent the well-known classes of starlike and convex functions, respectively. We denote $\mathcal{SS}^*(\gamma) := \mathcal{S}^*(((1+z)/(1-z))^\gamma)$ and $\mathcal{SK}(\gamma) := \mathcal{K}(((1+z)/(1-z))^\gamma)$ representing the class of strongly starlike and strongly convex functions of order $\gamma \in (0,1]$, respectively.

Recall that for two subfamilies G_1 and G_2 of $\mathcal{A}$, we say that r_0 is the G_1-radius for the class G_2 if r_0 $(0 < r \leq r_0)$ is the greatest number which satisfies $r^{-1}g(rz) \in G_1$, where $g \in G_2$. Moreover, starlike classes $\mathcal{S}^*(\psi)$ for different $\psi(z)$ were considered by many authors, whose works examined the geometrical properties, radius results and coefficient estimates of the functions of their respective classes. Sokół and Stankiewicz [20, 21] considered the class $\mathcal{S}_L^* := \mathcal{S}^*(\sqrt{1+z})$ and Mendiratta *et al.* [13] worked on the class $\mathcal{S}_{RL}^* := \mathcal{S}^*(\sqrt{2} - (\sqrt{2}-1)((1-z)/((1+2(\sqrt{2}-1)z)))^{1/2})$. Sharma *et al.* [19] studied the class $\mathcal{S}_C^* := \mathcal{S}^*(1+4z/3+2z^2/3)$ while the class $\mathcal{S}_s^* := \mathcal{S}^*(1+\sin z)$ was examined by Cho *et al.* [6]. The classes $\mathcal{S}_e^* := \mathcal{S}^*(e^z)$ and $\Delta^* := \mathcal{S}^*(z+\sqrt{1+z^2})$ were considered by Mendiratta *et al.* [14] and Raina *et al.* [15], respectively. Kargar *et al.* [10] introduced and studied the class $\mathcal{BS}^*(\alpha) := \mathcal{S}^*(1+z/(1-\alpha z^2))$, $\alpha \in [0,1]$, associated with the Booth lemniscate which was also investigated by Cho *et al.* [4]. Some more recent work on radius problems can be found in [1, 3, 5, 7, 23].

Motivated by the classes defined in [6, 10, 13–15, 19, 21], we consider the petal shaped region $\Omega_\rho := \{w \in \mathbb{C} : |\sinh(w-1)| < 1\}$, which is characterised functionally as $\rho(z) = 1 + \sinh^{-1}(z)$ to define our class. Clearly, $\rho(z)$ is a Ma-Minda function. See Figure 2 for its boundary curve γ_0 which is petal shaped. Note that $\sinh^{-1}(z)$ is a multivalued function and has the branch cuts along the line segments $(-i\infty, -i) \cup (i, i\infty)$, on the imaginary axis and hence it is analytic in $\mathbb{D}$. Now we introduce a new class of starlike functions

$$\mathcal{S}_\rho^* := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} \prec 1 + \sinh^{-1}(z) \right\} \quad (z \in \mathbb{D}), \tag{4}$$

which is associated with the petal-shaped domain $\rho(\mathbb{D})$. From the above definition, we deduce that $f \in \mathcal{S}_\rho^*$ if and only if there exists an analytic function $q(z) \prec \rho(z)$ such that

$$f(z) = z \exp\left(\int_0^z \frac{q(t)-1}{t} dt \right). \tag{5}$$

Table 1 presents some functions in the class $\mathcal{S}_\rho^*$, where $q_j \prec \rho$.

Since ρ is univalent in $\mathbb{D}$, $q_j(\mathbb{D}) \subset \rho(\mathbb{D})$ and $q_j(0) = \rho(0)$ $(j = 1, 2, 3)$, it follows that each $q_j \prec \rho$. Thus the functions $f_j(z)$ obtained from (5) are in the class $\mathcal{S}_\rho^*$. In particular, if we choose

$$q(z) = 1 + \sinh^{-1}(z) = 1 + z - \frac{z^3}{6} + \frac{3z^5}{40} - \frac{5z^7}{112} + \cdots,$$

TABLE 1. Some functions in the class $\mathcal{S}_\rho^*$

j	$q_j(z)$	$f_j(z)$
1	$1+z/5$	$z\exp(z/5)$
2	$(5+2z)/(5+z)$	$z+z^2/5$
3	$(7+4z)/(7+z)$	$z(1+z/7)^3$

then (5) gives

$$f_0(z) = z\exp\left(\int_0^z \frac{\sinh^{-1}(t)}{t}dt\right) = z + z^2 + \frac{z^3}{2} + \frac{z^4}{9} - \frac{z^5}{72} - \frac{z^6}{225} + \cdots, \tag{6}$$

which often acts as the extremal function for the class $\mathcal{S}_\rho^*$ yielding sharp results.

Remark 1.1. Note that $\sinh^{-1}(z) = \ln(z+\sqrt{1+z^2})$. Let $w = zf'(z)/f(z)$, where $f \in \mathcal{S}_\rho^*$. Then the class $\mathcal{S}_\rho^*$ can be alternatively represented by $\exp(w-1) \prec z+\sqrt{1+z^2}$, where $z+\sqrt{1+z^2}$ represents the Crescent shaped domain [15]. Thus, there exists an exponential relation among the functions in the classes $\mathcal{S}_\rho^*$ and Δ^*.

In the present investigation, the geometrical properties of the function $1+\sinh^{-1}(z)$ are studied and certain inclusion properties as well as radius problems are established for the class $\mathcal{S}_\rho^*$.

2. Properties of the function $1+\sinh^{-1}(z)$

The current section deals with the study of some geometric properties of the function $1+\sinh^{-1}(z)$.

Theorem 2.1. *The function $\rho(z) = 1+\sinh^{-1}(z)$ is a convex univalent function.*

Proof. Let $h(z) = \sinh^{-1}(z)$. Clearly, $h(0) = 0$. Since $h'(z) = 1/\sqrt{1+z^2}$ and $\sqrt{1+z^2} \prec \sqrt{1+z} \in \mathcal{P}$, where $\mathcal{P}$ is the Carathéodory class. Therefore, $1/\sqrt{1+z^2} \in \mathcal{P}$ which implies that $\operatorname{Re} h'(z) > 0$. Hence ρ is univalent. Now a calculation yields

$$1 + \frac{zh''(z)}{h'(z)} = \frac{1}{1+z^2}.$$

Since

$$\frac{1}{1+z^2} \prec \frac{1}{1+z} \in \mathcal{P}.$$

Therefore, $\operatorname{Re}(1+zh''(z)/h'(z)) > 0$ which implies that h (and thus ρ) is a convex univalent function. □

Remark 2.2. Note that $\rho'(0) > 0$ and the function $\varphi(z) = z+\sqrt{1+z^2}$ satisfies $\varphi(\bar{z}) = \overline{\varphi(z)}$. Therefore, $\rho(\bar{z}) = \overline{\rho(z)}$ and hence, the domain $\Omega_\rho = \rho(\mathbb{D})$ is symmetric about the real axis.

Theorem 2.3. *The domain Ω_ρ is symmetric about the line* $\mathrm{Re}(w) = 1$.

Proof. Since Ω_ρ is symmetric about the real axis, the condition $0 \le \theta \le \pi/2$ is sufficient to prove our result. As we know that symmetry along imaginary axis for $f \in \mathcal{A}$ holds if $\mathrm{Re}(f(\theta)) = -\mathrm{Re}(f(\pi-\theta))$ and $\mathrm{Im}(f(\theta)) = \mathrm{Im}(f(\pi-\theta))$. Now let $h(z) = \sinh^{-1}(z) = \ln(z + \sqrt{1+z^2})$. Then $\mathrm{Im}(h(z)) = \arg(z + \sqrt{1+z^2})$. For $z = re^{it}$, $t \in [0, \pi]$ and fixed $r \in (0, 1)$, we have the following expressions for $t \to \theta$

$$\begin{aligned} I_1 &= \arg\left(r(\cos\theta + i\sin\theta) + \sqrt{1 + r^2(\cos(2\theta) + i\sin(2\theta))}\right) \\ &= \arg\left(z + \sqrt{1+z^2}\right), \end{aligned}$$

and for $t \to \pi - \theta$

$$\begin{aligned} I_2 &= \arg\Big(r(\cos(\pi-\theta) + i\sin(\pi-\theta)) \\ &\qquad + \sqrt{1 + r^2(\cos(2(\pi-\theta)) + i\sin(2(\pi-\theta)))}\Big) \\ &= \arg\left(r(-\cos\theta + i\sin\theta) + \sqrt{1 + r^2(\cos(2\theta) - i\sin(2\theta))}\right) \\ &= \arg\left(-\overline{z} + \sqrt{1+\overline{z}^2}\right). \end{aligned}$$

Now let us consider $(z + \sqrt{1+z^2})/(-\overline{z} + \sqrt{1+\overline{z}^2})$. On rationalising the denominator, we get

$$\frac{z + \sqrt{1+z^2}}{-\overline{z} + \sqrt{1+\overline{z}^2}} = \frac{(z + \sqrt{1+z^2})(-z + \sqrt{1+z^2})}{(-\overline{z} + \sqrt{1+\overline{z}^2})(-z + \sqrt{1+z^2})} = \frac{1}{|-z + \sqrt{1+z^2}|^2} = k,$$

where $k > 0$ is some real positive constant. Thus,

$$\begin{aligned} & \arg\left(\frac{z + \sqrt{1+z^2}}{-\overline{z} + \sqrt{1+\overline{z}^2}}\right) = \arg(k) = 0 \\ \Rightarrow\quad & \arg\left(z + \sqrt{1+z^2}\right) = \arg\left(-\overline{z} + \sqrt{1+\overline{z}^2}\right) \\ \Rightarrow\quad & I_1 = I_2. \end{aligned}$$

Similarly, $\mathrm{Re}(h(\theta)) = -\mathrm{Re}(h(\pi-\theta))$ for $0 \le \theta \le \pi/2$. Hence, $h(z)$ is symmetric about the imaginary axis and thus, by translation property, $\rho(z)$ is symmetric about the line $\mathrm{Re}(w) = 1$. □

Now using Theorem 2.3, we obtain the next result:

Corollary 2.4. *The disk* $\{w : |w-1| \le \sinh^{-1}(r)\}$ *is contained in* $\rho(|z| \le r)$ *and is maximal.*

Proof. Since $\min_{|z|=r} |\sinh^{-1}(z)| = |\sinh^{-1}(-r)| = \sinh^{-1}(r)$ and hence the conclusion can be drawn at once. □

Theorem 2.5. *We find that the following properties hold for* $\rho(z) = 1 + \sinh^{-1}(z)$:

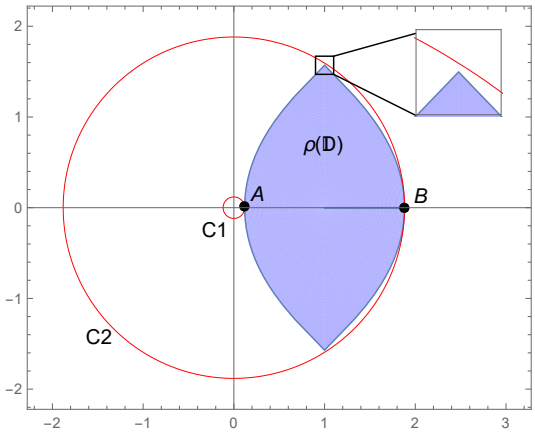


FIGURE 1. $\rho(\mathbb{D})$ lies in the annular region bounded between the circles $C1$ and $C2$.

(i) $\rho(-r) \le \operatorname{Re}\rho(z) \le \rho(r) \quad (|z| \le r < 1)$;
(ii) $|\operatorname{Im}\rho(z)| \le \pi/2 \quad (|z| \le 1)$;
(iii) $\rho(-r) \le |\rho(z)| \le \rho(r) \quad (|z| \le r < 1)$;
(iv) $|\arg\rho(z)| \le \tan^{-1}(1/t)$, *where* $t = \frac{4}{\pi}\sqrt{\sinh^{-1}(1)(1-\sinh^{-1}(1))}$.

Proof. (i) Since $\rho(z)$ is convex and typically real, the value of $\operatorname{Re}\rho(z)$ falls between $\lim_{\theta\to 0}\rho(re^{\theta})$ and $\lim_{\theta\to\pi}\rho(re^{\theta})$, thus the result follows.

(ii) Using Theorem 2.3, it suffices to take $\theta \in [0, \pi/2]$. Then the inequality follows by letting r tending to 1^- and observing that the function

$$\operatorname{Im}\rho(z) = \arg\left(r\cos(\theta) + \sqrt{1 + r^2(\cos(2\theta) + i\sin(2\theta))} + ir\sin(\theta)\right)$$

is strictly increasing in the interval $[0, \pi/2]$ and hence the result follows at once.

(iii) The radially farthest and nearest points in $\rho(\mathbb{D})$ from origin are respectively B and A (see Figure 1) and therefore the result obviously holds. Moreover we observe that these points A and B lie on the real line and hence the bounds of $|\rho(z)|$ and $\operatorname{Re}\rho(z)$ coincide.

The proof of (iv) is evident from Theorem 3.1(iii) so skipped here. □

Next we have the following important result:

Lemma 2.6. *For* $1 - \sinh^{-1}(1) < a < 1 + \sinh^{-1}(1)$, *let* r_a *be given by*

$$r_a = \begin{cases} a - (1 - \sinh^{-1}(1)), & 1 - \sinh^{-1}(1) < a \le 1; \\ 1 + \sinh^{-1}(1) - a, & 1 \le a < 1 + \sinh^{-1}(1). \end{cases}$$

Then $\{w : |w - a| < r_a\} \subset \Omega_{\rho}$.

We omit the proof of Lemma 2.6 as it directly follows from Theorem 2.3 and Corollary 2.4.

Remark 2.7. Evidently the domain Ω_ρ is contained inside the disk $\{w : |w-1| < \pi/2\}$.

3. Inclusion relations

This section establishes some inclusion results involving the class $\mathcal{S}^*_\rho$ with some well-known classes.

We consider the class $M(\beta)$, first studied by Uralegaddi *et al.* [22], given by

$$M(\beta) := \left\{ f \in \mathcal{A} : \operatorname{Re}\left(\frac{zf'(z)}{f(z)}\right) < \beta,\ z \in \mathbb{D},\ \beta > 1 \right\},$$

and another interesting class introduced by Kanas and Wisńiowska [9] of k-starlike functions, denoted by $k - \mathcal{ST}$ and defined by

$$k - \mathcal{ST} := \left\{ f \in \mathcal{A} : \operatorname{Re}\left(\frac{zf'(z)}{f(z)}\right) > k \left|\frac{zf'(z)}{f(z)} - 1\right|,\ z \in \mathbb{D},\ k \geq 0 \right\}.$$

Note that $\mathcal{S}^* = 0 - \mathcal{ST}$ and $\mathcal{S}^*_p = 1 - \mathcal{ST}$, where $\mathcal{S}^*_p$ is the class of parabolic starlike functions [17].

We establish the following inclusion relations for the class $\mathcal{S}^*_\rho$.

Theorem 3.1. *The class $\mathcal{S}^*_\rho$ satisfies the following relationships*:

(i) $\mathcal{S}^*_\rho \subset \mathcal{S}^*_\alpha \subset \mathcal{S}^*$ *for* $0 \leq \alpha \leq 1 - \sinh^{-1}(1)$;
(ii) $\mathcal{S}^*_\rho \subset M(\beta)$ *for* $\beta \geq 1 + \sinh^{-1}(1)$;
(iii) $\mathcal{S}^*_\rho \subset \mathcal{SS}^*(\gamma)$ *for* $(2/\pi)\tan^{-1}(1/t) \leq \gamma \leq 1$,
where $t = \frac{4}{\pi}\sqrt{\sinh^{-1}(1)(1 - \sinh^{-1}(1))}$;
(iv) $k - \mathcal{ST} \subset \mathcal{S}^*_\rho$ *for* $k \geq 1 + 1/\sinh^{-1}(1)$.

Proof. Consider $f \in \mathcal{S}^*_\rho$ which implies $zf'(z)/f(z) \prec 1 + \sinh^{-1}(z)$. By Theorem 2.5, it is evident that for $z \in \mathbb{D}$,

$$1 - \sinh^{-1}(1) = \min_{|z|=1} \operatorname{Re}(1 + \sinh^{-1}(z)) \leq \operatorname{Re}\frac{zf'(z)}{f(z)}$$

and

$$\operatorname{Re}\frac{zf'(z)}{f(z)} \leq \max_{|z|=1} \operatorname{Re}(1 + \sinh^{-1}(z)) = 1 + \sinh^{-1}(1).$$

This proves (i) and (ii).

For (iii), let $w \in \mathbb{C}$, $X = \operatorname{Re}(w)$, $Y = \operatorname{Im}(w)$, and $b = 1 - \sinh^{-1}(1)$. Now consider the parabolic domain Γ_P with the boundary curve $\partial\Gamma_P = \gamma_p : Y^2 = 4a(X - b)$. Then the focus a of the smallest parabola γ_p which contains Ω_ρ will touch the peak points $1 \pm i\pi/2$ of $\mathcal{S}^*_\rho$ is $\pi^2/(16\sinh^{-1}(1))$. Let P be any point on the parabola γ_P with parametric coordinates $(b + at^2, 2at)$ such that the tangent OE at P passes through origin for some parameter t. Let the equation

of the tangent OE be $y = mx$, where $m = dy/dx = (dy/dt)/(dx/dt) = 1/t$. Therefore at P, we have

$$m = \frac{y}{x} \Rightarrow \frac{1}{t} = \frac{2at}{b + at^2},$$

which yields

$$t = \sqrt{\frac{b}{a}} = \frac{4}{\pi}\sqrt{\sinh^{-1}(1)(1 - \sinh^{-1}(1))} \tag{7}$$

and the argument of the tangent at P of γ_p is $\tan^{-1}(1/t)$. Since $\Omega_\rho \subset \Gamma_p$, it gives

$$\left|\arg \frac{zf'(z)}{f(z)}\right| \leq \max_{|z|=1} \arg(\rho(z)) = \max \arg(\gamma_p) = \tan^{-1}(1/t),$$

which demonstrates $f \in \mathcal{SS}^*((2/\pi)\tan^{-1}(1/t))$, where t is given by (7).

To show (iv), consider $f \in k - \mathcal{ST}$ along with the conic domain $\Gamma_k = \{w \in \mathbb{C} : \operatorname{Re} w > k|w - 1|\}$. For $k > 1$, let $\partial\Gamma_k$ represent the horizontal ellipse $\gamma_k : x^2 = k^2(x-1)^2 + k^2y^2$ which may be rewritten as

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1,$$

where $x_0 = k^2/(k^2 - 1)$, $y_0 = 0$, $a = k/(k^2 - 1)$ and $b = 1/\sqrt{k^2 - 1}$. For $\gamma_k \subset \Omega_\rho$, the condition $x_0 + a \leq 1 + \sinh^{-1}(1)$ must hold, or equivalently $k \geq 1 + 1/\sinh^{-1}(1)$. Since $\Gamma_{k_1} \subseteq \Gamma_{k_2}$ for $k_1 \geq k_2$, it follows that for $k \geq 1 + 1/\sinh^{-1}(1)$, $k - \mathcal{ST} \subset \mathcal{S}^*_\rho$. Figure 2 clearly depicts these relations. □

For our next result, we consider $\mathcal{P}_n[C, D]$, the class of functions $p(z)$ of the form $1 + \sum_{k=n}^{\infty} c_k z^k$, satisfying $p(z) \prec (1 + Cz)/(1 + Dz)$, where $-1 \leq D < C \leq 1$. Denote by $\mathcal{P}_n(\alpha) := \mathcal{P}_n[1 - 2\alpha, -1]$ and $\mathcal{P}_n := \mathcal{P}_n(0)$. For $n = 1$, $\mathcal{P} = \mathcal{P}_1$ is the Carathéodory class. We need the following lemmas:

Lemma 3.2 ([18]). *For $p \in \mathcal{P}_n(\alpha)$, we have*

$$\left|\frac{zp'(z)}{p(z)}\right| \leq \frac{2(1-\alpha)nr^n}{(1 - r^n)(1 + (1 - 2\alpha)r^n)}, \ (|z| = r).$$

Lemma 3.3 ([16]). *For $p \in \mathcal{P}_n[C, D]$, we have*

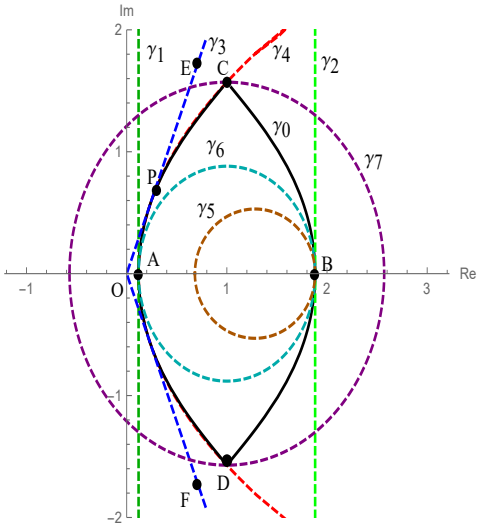
$$\left|p(z) - \frac{1 - CDr^{2n}}{1 - D^2r^{2n}}\right| \leq \frac{(C - D)r^n}{1 - D^2r^{2n}}, \ (|z| = r).$$

Especially, for $p \in \mathcal{P}_n(\alpha)$, we have

$$\left|p(z) - \frac{1 + (1 - 2\alpha)r^{2n}}{1 - r^{2n}}\right| \leq \frac{2(1-\alpha)r^n}{1 - r^{2n}}, \ (|z| = r).$$

Theorem 3.4. *Let $-1 < D < C \leq 1$. If either of the following two conditions holds*:

(i) $(1-\sinh^{-1}(1))(1-D^2) < 1-CD \leq 1-D^2$ *and* $C-D \leq (1-D)\sinh^{-1}(1)$;
(ii) $1-D^2 \leq 1-CD < (1+\sinh^{-1}(1))(1-D^2)$ *and* $C-D \leq (1+D)\sinh^{-1}(1)$.



$\gamma_0 : w = 1 + \sinh^{-1}(z)$
$\gamma_1 : \mathrm{Re}(w) = 1 - \sinh^{-1}(1)$
$\gamma_2 : \mathrm{Re}(w) = 1 + \sinh^{-1}(1)$
$\gamma_3 : |\arg(w)| \leq \tan^{-1}(1/t)$
where t is given by (7)
$\gamma_4 = \gamma_p : Y^2 = 4a(X - 1 + \sinh^{-1}(1))$
where $a = \pi^2/(16 \sinh^{-1}(1))$
$\gamma_5 = \gamma_k : \mathrm{Re}(w) > k|w - 1|$
where $k = 1 + 1/\sinh^{-1}(1)$
$\gamma_6 = D_L : |w - 1| < \sinh^{-1}(1)$
$\gamma_7 = D_S : |w - 1| < \pi/2$
$\mathrm{Re(A)} = 1 - \sinh^{-1}(1)$
$\mathrm{Re(B)} = 1 + \sinh^{-1}(1)$
$\mathrm{Im(C)} = -\mathrm{Im}(D) = \pi/2$
$\arg(\mathrm{E}) = -\arg(\mathrm{F}) = \tan^{-1}(1/t)$
where t is given by (7)
O: origin
P: point of tangency of γ_4 w.r.t. OE

FIGURE 2. Boundary curves, depicting some inclusion relations for $w = 1 + \sinh^{-1}(z)$.

Then $\mathcal{S}^*[C, D] \subset \mathcal{S}^*_\rho$.

Proof. Let $f \in \mathcal{S}^*[C, D]$ which implies $zf'(z)/f(z) \in \mathcal{P}[C, D]$. Using Lemma 3.3 we have

$$\left| \frac{zf'(z)}{f(z)} - \frac{1 - CD}{1 - D^2} \right| \leq \frac{(C - D)}{1 - D^2}. \tag{8}$$

Let $a = (1 - CD)/(1 - D^2)$ and assume that (i) holds. Now multiplying $1 + D$ and dividing by $(1 - D^2)$ on either side of the inequality $(C - D) \leq (1 - D)\sinh^{-1}(1)$ gives $(C - D)/(1 - D^2) \leq a - (1 - \sinh^{-1}(1))$ on simplification. Also, the inequality $(1 - \sinh^{-1}(1))(1 - D^2) < 1 - CD \leq 1 - D^2$ is equivalent to $1 - \sinh^{-1}(1) < (1 - CD)/(1 - D^2) \leq 1$. Therefore, from (8) we find $w = zf'(z)/f(z)$ is contained inside the disk $|w - a| < r_a$, where $r_a = a - (1 - \sinh^{-1}(1))$ and $1 - \sinh^{-1}(1) < a \leq 1$. Hence $f \in \mathcal{S}^*_\rho$ by Lemma 2.6. A similar proof can be shown when (ii) holds. □

4. Radius problems

In this section, radius results for various subclasses of $\mathcal{A}$ are established. We begin by determining sharp $\mathcal{S}^*_\alpha$ $(0 \leq \alpha < 1)$, $\mathcal{M}(\beta)$ $(\beta > 1)$ and $k - \mathcal{ST}$-radii $(k \geq 0)$ for the class $\mathcal{S}^*_\rho$. Using Theorem 3.1, we can establish that $R_{\mathcal{S}^*_\alpha}(\mathcal{S}^*_\rho) = R_{M(\beta)}(\mathcal{S}^*_\rho) = 1$ for $0 \leq \alpha \leq 1 - \sinh^{-1}(1)$ and $\beta > 1 + \sinh^{-1}(1)$.

Theorem 4.1. *If* $f \in \mathcal{S}^*_\rho$, *then the following results hold*:

(i) *For* $1-\sinh^{-1}(1)\le\alpha<1$, *we have* $f\in\mathcal{S}^*_\alpha$ *in* $|z|\le\sinh(1-\alpha)$.
(ii) *For* $1<\beta\le 1+\sinh^{-1}(1)$, *we have* $f\in\mathcal{M}(\beta)$ *in* $|z|\le\sinh(\beta-1)$.
(iii) *For* $k>0$, *we have* $f\in k-\mathcal{ST}$ *in* $|z|\le\sinh(1/(k+1))$.

The results are sharp.

Proof. Since $f\in\mathcal{S}^*_\rho$, $zf'(z)/f(z)\prec 1+\sinh^{-1}(z)$ and hence for $|z|=r<1$ Theorem 2.5 gives

$$1-\sinh^{-1}(r)\le\operatorname{Re}\frac{zf'(z)}{f(z)}\le 1+\sinh^{-1}(r),$$

thereby validating the first two parts. Also, the constants $\sinh(1-\alpha)$ and $\sinh(\beta-1)$ are optimal for the function f_0 given by (6). Now to prove (iii), note that $f\in k-\mathcal{ST}$ in $|z|<r$, if

$$\operatorname{Re}(1+\sinh^{-1}(w(z)))\ge k|1+\sinh^{-1}(w(z))-1|=k|\sinh^{-1}(w(z))|.$$

Here w denotes the Schwarz function. Since $\operatorname{Re}(1+\sinh^{-1}(w(z)))\ge 1-\sinh^{-1}(r)$ and $|\sinh^{-1}(w(z))|\le\sinh^{-1}(r)$, the inequality $\operatorname{Re}(1+\sinh^{-1}(w(z)))\ge k|\sinh^{-1}(w(z))|$ holds whenever $1-\sinh^{-1}(r)\ge k\sinh^{-1}(r)$, which implies $r\le\sinh(1/(1+k))$. For the function f_0 given by (6) and for $z_0=-\sinh(1/(1+k))$, we have

$$\operatorname{Re}\frac{z_0f_0'(z_0)}{f_0(z_0)}=\operatorname{Re}(1+\sinh^{-1}(z_0))=\frac{k}{k+1}=k|\sinh^{-1}(z_0)|=k\left|\frac{z_0f_0'(z_0)}{f_0(z_0)}-1\right|.$$

This concludes the proof. □

Corollary 4.2. *Substituting* $k=1$ *in part* (iii) *above, we find that* $f\in\mathcal{S}^*_\rho$ *is parabolic starlike* [17] in $|z|\le\sinh(1/2)$.

In the next result, we find the $\mathcal{K}_\alpha$-radius for the class $\mathcal{S}^*_\rho$.

Theorem 4.3. *Let* $f\in\mathcal{S}^*_\rho$. *Then* $f\in\mathcal{K}_\alpha$ *in* $|z|<r_\alpha$, *where* r_α *is the least positive root of*

$$(1-r^2)\sqrt{1+r^2}\left(1-\sinh^{-1}(r)\right)\left(1-\alpha-\sinh^{-1}(r)\right)-r=0\quad(0\le\alpha<1).\tag{9}$$

Proof. Let $f\in\mathcal{S}^*_\rho$ and w be a Schwarz function. Then $zf'(z)/f(z)=1+\sinh^{-1}(w(z))$ such that

$$1+\frac{zf''(z)}{f'(z)}=1+\sinh^{-1}(w(z))+\frac{zw'(z)}{(1+\sinh^{-1}(w(z)))\sqrt{1+w^2(z)}}$$

which yields

$$\operatorname{Re}\left(1+\frac{zf''(z)}{f'(z)}\right)\ge\operatorname{Re}\left(1+\sinh^{-1}(w(z))\right)-\left|\frac{zw'(z)}{(1+\sinh^{-1}(w(z)))\sqrt{1+w^2(z)}}\right|.$$

We know for the Schwarz function w, the inequality $|w'(z)| \leq (1-|w(z)|^2)/(1-|z|^2)$ holds. Thus we observe that

$$\begin{aligned}\operatorname{Re}\left(1+\frac{zf''(z)}{f'(z)}\right) &\geq 1-\sinh^{-1}(|z|)-\frac{|z|(1-|w(z)|^2)}{(1-\sinh^{-1}(|z|))(1-|z|^2)\sqrt{1+|z|^2}}\\ &\geq 1-\sinh^{-1}(|z|)-\frac{|z|}{(1-\sinh^{-1}(|z|))(1-|z|^2)\sqrt{1+|z|^2}}.\end{aligned}$$

Let $q(r) := 1-\sinh^{-1}(r)-r/\left((1-\sinh^{-1}(r))(1-r^2)\sqrt{1+r^2}\right)$. We find $q(r)$ is a decreasing function in $[0,1)$ with $q(0)=1$. Therefore $\operatorname{Re}(1+zf''(z)/f'(z)) > \alpha$ in $|z| < r_\alpha < 1$, where r_α is given as the least positive root of the equation $q(r)=\alpha$, which is same as (9) and hence the result. □

Remark 4.4. Note for $\alpha = 0$, $r_0 \approx 0.37198$ which is not sharp, so the result can be further improved. The sharp $\mathcal{K}_0$-radius for the class $\mathcal{S}^*_\rho$ is $r_0 \approx 0.400435$, which we can guess graphically but a mathematical proof is yet to derive.

For our next Theorems 4.5–4.8, the following subclasses are required:

Let $\mathcal{S}^*_n[C,D] := \{f \in \mathcal{A}_n : zf'(z)/f(z) \in \mathcal{P}_n[C,D]\}$. Also, let $\mathcal{S}^*_n(\alpha) := \mathcal{S}^*_n[1-2\alpha,-1] = \mathcal{A}_n \cap \mathcal{S}^*_\alpha$ and $\mathcal{S}^*_{\rho,n} := \mathcal{A}_n \cap \mathcal{S}^*_\rho$. Further, Ali *et al.* [2] studied the three classes $\mathcal{S}_n := \{f \in \mathcal{A}_n : f(z)/z \in \mathcal{P}_n\}$, $\mathcal{S}^*_n[C,D]$ and

$$\mathcal{CS}_n(\alpha) := \left\{ f \in \mathcal{A}_n : \frac{f(z)}{g(z)} \in \mathcal{P}_n,\ g \in \mathcal{S}^*_n(\alpha) \right\}.$$

Now we obtain the $\mathcal{S}^*_{\rho,n}$-radii for the classes defined above.

Theorem 4.5. *For the class $\mathcal{S}_n$, the sharp $\mathcal{S}^*_{\rho,n}$-radius is given by*:

$$R_{\mathcal{S}^*_{\rho,n}}(\mathcal{S}_n) = \left(\frac{\sinh^{-1}(1)}{n+\sqrt{n^2+\left(\sinh^{-1}(1)\right)^2}}\right)^{1/n}.$$

Proof. Let $f \in \mathcal{S}_n$. Define $s : \mathbb{D} \to \mathbb{C}$ by $s(z) = f(z)/z$. Then $s \in \mathcal{P}_n$ and we can obtain $zf'(z)/f(z) - 1 = zs'(z)/s(z)$ from the above definition of s. Using Lemma 2.6 and Lemma 3.2, the following holds

$$\left|\frac{zf'(z)}{f(z)}-1\right| = \left|\frac{zs'(z)}{s(z)}\right| \leq \frac{2nr^n}{1-r^{2n}} \leq \sinh^{-1}(1),$$

or equivalently $(\sinh^{-1}(1))r^{2n} + 2nr^n - \sinh^{-1}(1) \leq 0$. Therefore, the $\mathcal{S}^*_{\rho,n}$-radius of $\mathcal{S}_n$ is the least positive root of $(\sinh^{-1}(1))r^{2n} + 2nr^n - \sinh^{-1}(1) = 0$ for $r \in (0,1)$. We can verify $\operatorname{Re}(f_0(z)/z) > 0$ holds in $\mathbb{D}$, where $f_0(z) = z(1+z^n)/(1-z^n)$. Thus $f_0 \in \mathcal{S}_n$ and $zf_0'(z)/f_0(z) = 1 + 2nz^n/(1-z^{2n})$. Moreover, the result is sharp since at $z = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{S}_n)$, we obtain

$$\frac{zf_0'(z)}{f_0(z)} - 1 = \frac{2nz^n}{1-z^{2n}} = \sinh^{-1}(1).$$

The proof is complete. □

Let $\mathcal{F}$ define the class of functions $f \in \mathcal{A}$ satisfying $f(z)/z \in \mathcal{P}$. The radius of univalence and starlikeness of the class $\mathcal{F}$ is $\sqrt{2}-1$, as shown in [12].

Corollary 4.6. *For the class $\mathcal{F}$, the $\mathcal{S}^*_\rho$-radius is stated as*

$$R_{\mathcal{S}^*_\rho}(\mathcal{F}) = -e + \sqrt{1+e^2} \approx 0.178105.$$

Theorem 4.7. *For the class $\mathcal{CS}_n(\alpha)$, the sharp $\mathcal{S}^*_{\rho,n}$-radius is given by*

$$R_{\mathcal{S}^*_{\rho,n}}(\mathcal{CS}_n(\alpha)) = \left(\frac{\sinh^{-1}(1)}{n-\alpha+1+\sqrt{(n-\alpha+1)^2+(\sinh^{-1}(1)+2(1-\alpha))\sinh^{-1}(1)}} \right)^{1/n}.$$

Proof. Let $f \in \mathcal{CS}_n(\alpha)$ and $g \in \mathcal{S}^*_n(\alpha)$. Considering $s(z) = f(z)/g(z)$, clearly indicates $s \in \mathcal{P}_n$. Also, it gives

$$\frac{zf'(z)}{f(z)} = \frac{zs'(z)}{s(z)} + \frac{zg'(z)}{g(z)}.$$

The use of Lemmas 3.2–3.3 gives us

$$\left| \frac{zf'(z)}{f(z)} - \frac{1+(1-2\alpha)r^{2n}}{1-r^{2n}} \right| \leq \frac{2(n-\alpha+1)r^n}{1-r^{2n}}. \tag{10}$$

Considering $(1+(1-2\alpha)r^{2n})/(1-r^{2n}) \geq 1$, the relation $f \in \mathcal{S}^*_{\rho,n}$ follows from (10) and Lemma 2.6 if the subsequent inequality is true:

$$\frac{1+2(n-\alpha+1)r^n+(1-2\alpha)r^{2n}}{1-r^{2n}} \leq 1+\sinh^{-1}(1)$$

or equivalently, $(2-2\alpha+\sinh^{-1}(1))r^{2n}+2(n-\alpha+1)r^n-\sinh^{-1}(1) \leq 0$ holds. Thus, the least positive root of

$$(2-2\alpha+\sinh^{-1}(1))r^{2n}+2(n-\alpha+1)r^n-\sinh^{-1}(1) = 0$$

gives the $\mathcal{S}^*_{\rho,n}$-radius for the class $\mathcal{CS}_n(\alpha)$. Next examine the following functions

$$f_0(z) = \frac{z(1+z^n)}{(1-z^n)^{(n+2-2\alpha)/n}} \text{ and } g_0(z) = \frac{z}{(1-z^n)^{2(1-\alpha)/n}}, \tag{11}$$

which implies $f_0(z)/g_0(z) = (1+z^n)/(1-z^n)$ and $zg_0'(z)/g_0(z) = (1+(1-2\alpha)z^n)/(1-z^n)$. Moreover, $\mathrm{Re}(f_0(z)/g_0(z)) > 0$ and $\mathrm{Re}(zg_0'(z)/g_0(z)) > \alpha$ in the unit disk $\mathbb{D}$ is obvious. Hence $f_0 \in \mathcal{CS}_n(\alpha)$. At $z = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{CS}_n(\alpha))$, the function f_0 defined in (11) satisfies

$$\frac{zf_0'(z)}{f_0(z)} = \frac{1+2(n-\alpha+1)z^n+(1-2\alpha)z^{2n}}{1-z^{2n}} = 1+\sinh^{-1}(1),$$

which accomplish sharpness of the result. □

Theorem 4.8. *For the class $\mathcal{S}^*_n[C,D]$, the $\mathcal{S}^*_{\rho,n}$-radius is given by*

$$R_{\mathcal{S}^*_{\rho,n}}(\mathcal{S}^*_n[C,D]) = \begin{cases} \min\{1;R_1\}, & -1 \leq D < 0 < C \leq 1; \\ \min\{1;R_2\}, & 0 < D < C \leq 1, \end{cases}$$

where

$$R_1 := \left(\frac{2\sinh^{-1}(1)}{C - D + \sqrt{(C-D)^2 + 4(D^2(1+\sinh^{-1}(1)) - CD)\sinh^{-1}(1)}} \right)^{\frac{1}{n}}$$

and

$$R_2 := \left(\frac{2\sinh^{-1}(1)}{C - D + \sqrt{(C-D)^2 + 4(D^2(\sinh^{-1}(1)-1) + CD)\sinh^{-1}(1)}} \right)^{\frac{1}{n}}.$$

Proof. Let $f \in \mathcal{S}_n^*[C, D]$. From Lemma 3.3, we have

$$\left| \frac{zf'(z)}{f(z)} - b \right| \leq \frac{(C-D)r^n}{1 - D^2 r^{2n}}, \tag{12}$$

where $b = (1 - CDr^{2n})/(1 - D^2 r^{2n})$, $|z| = r$, represents the center of the disk. We infer $b \geq 1$ for $-1 \leq D < 0 < C \leq 1$. From Lemma 2.6, $f \in \mathcal{S}_{\rho,n}^*$ depends on whether following condition is true:

$$\frac{1 + (C-D)r^n - CDr^{2n}}{1 - D^2 r^{2n}} \leq 1 + \sinh^{-1}(1),$$

which reduces to

$$r \leq \left(\frac{2\sinh^{-1}(1)}{C - D + \sqrt{(C-D)^2 + 4(D^2(1+\sinh^{-1}(1)) - CD)\sinh^{-1}(1)}} \right)^{\frac{1}{n}} = R_1.$$

Further, taking $D = 0$, we get $b = 1$. Then (12) yields

$$\left| \frac{zf'(z)}{f(z)} - 1 \right| \leq Cr^n, \ (0 < C \leq 1).$$

Now applying Lemma 2.6 with $a = 1$ gives $f \in \mathcal{S}_{\rho,n}^*$ if $r \leq ((\sinh^{-1}(1))/C)^{1/n}$.

For $0 < D < C \leq 1$, we have $b < 1$. Thus, using Lemma 2.6 and (12), we have $f \in \mathcal{S}_{\rho,n}^*$ if the following holds:

$$\frac{CDr^{2n} + (C-D)r^n - 1}{1 - D^2 r^{2n}} \leq \sinh^{-1}(1) - 1,$$

or equivalently, if

$$r \leq \left(\frac{2\sinh^{-1}(1)}{C - D + \sqrt{(C-D)^2 + 4(D^2(\sinh^{-1}(1)-1) + CD)\sinh^{-1}(1)}} \right)^{\frac{1}{n}} = R_2.$$

This concludes the proof. □

The next theorem establishes radius results for some well-known classes mentioned earlier.

Theorem 4.9. *The sharp $\mathcal{S}_\rho^*$-radii for the classes $\mathcal{S}_L^*$, $\mathcal{S}_{RL}^*$, $\mathcal{S}_C^*$, $\mathcal{S}_e^*$, Δ^* and $\mathcal{BS}^*(\alpha)$ are*:

(i) $R_{\mathcal{S}_\rho^*}(\mathcal{S}_L^*) = \sinh^{-1}(1)(2-\sinh^{-1}(1)) \approx 0.985928.$

(ii) $R_{\mathcal{S}_\rho^*}(\mathcal{S}_{RL}^*) = \frac{\left(2+(1+\sqrt{2})\sinh^{-1}(1)\right)\sinh^{-1}(1)}{5-3\sqrt{2}+\left(4(\sqrt{2}-1)+2\sinh^{-1}(1)\right)\sinh^{-1}(1)} \approx 0.964694.$

(iii) $R_{\mathcal{S}_\rho^*}(\mathcal{S}_C^*) = \frac{1}{2}\left(\sqrt{2\left(2+3\sinh^{-1}(1)\right)}-2\right) \approx 0.523831.$

(iv) $R_{\mathcal{S}_\rho^*}(\mathcal{S}_e^*) = \ln(1+\sinh^{-1}(1)) \approx 0.632002.$

(v) $R_{\mathcal{S}_\rho^*}(\Delta^*) = \frac{\sinh^{-1}(1)(2+\sinh^{-1}(1))}{2(1+\sinh^{-1}(1))} \approx 0.674924.$

(vi) $R_{\mathcal{S}_\rho^*}(\mathcal{BS}^*(\alpha)) = \frac{-1+\sqrt{1+\alpha\left(2\sinh^{-1}(1)\right)^2}}{2\alpha\sinh^{-1}(1)},\ \alpha \in [0,1].$

Proof. (i) Suppose $f \in \mathcal{S}_L^*$. We have $zf'(z)/f(z) \prec \sqrt{1+z}$. When $|z| = r$, we obtain

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq 1 - \sqrt{1-r} \leq \sinh^{-1}(1),$$

such that $r \leq (2-\sinh^{-1}(1))\sinh^{-1}(1) = R_{\mathcal{S}_\rho^*}(\mathcal{S}_L^*)$ holds. Next examine the function

$$f_0(z) = \frac{4z}{\left(1+\sqrt{1+z}\right)^2} e^{2\left(\sqrt{1+z}-1\right)}.$$

Since $zf_0'(z)/f_0(z) = \sqrt{1+z}$, it follows that $f_0 \in \mathcal{S}_L^*$. As $zf_0'(z)/f_0(z) - 1 = -\sinh^{-1}(1)$ is obtained at $z = -R_{\mathcal{S}_\rho^*}(\mathcal{S}_L^*)$, the result is sharp.

(ii) Suppose $f \in \mathcal{S}_{RL}^*$, we obtain

$$\frac{zf'(z)}{f(z)} \prec \sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1-z}{(1+2(\sqrt{2}-1)z)}}.$$

For $|z| = r$, the subsequent inequality holds

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq 1 - \sqrt{2} + (\sqrt{2}-1)\sqrt{\frac{1+r}{(1-2(\sqrt{2}-1)r)}} \leq \sinh^{-1}(1),$$

provided

$$r \leq \frac{\left(2+(1+\sqrt{2})\sinh^{-1}(1)\right)\sinh^{-1}(1)}{5-3\sqrt{2}+\left(4(\sqrt{2}-1)+2\sinh^{-1}(1)\right)\sinh^{-1}(1)} = R_{\mathcal{S}_\rho^*}(\mathcal{S}_{RL}^*).$$

Next observe the following function defined as

$$f_0(z) = z\exp\left(\int_0^z \frac{q_0(t)-1}{t}dt\right),$$

where

$$q_0(t) = \sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1-t}{(1+2(\sqrt{2}-1)t)}}.$$

From the definition of f_0, at $z = -R_{\mathcal{S}^*_\rho}(\mathcal{S}^*_{RL})$, we have

$$\frac{zf_0'(z)}{f_0(z)} = \sqrt{2} - (\sqrt{2}-1)\sqrt{\frac{1-z}{(1+2(\sqrt{2}-1)z)}} = 1 - \sinh^{-1}(1),$$

which confirms the sharpness.

(iii) Suppose $f \in \mathcal{S}^*_C$. So $zf'(z)/f(z) \prec 1 + 4z/3 + 2z^2/3$. This gives

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{4r}{3} + \frac{2r^2}{3} \leq \sinh^{-1}(1),\ |z| = r,$$

for $r \leq \frac{1}{2}\left(\sqrt{2\left(2+3\sinh^{-1}(1)\right)} - 2\right) = R_{\mathcal{S}^*_\rho}(\mathcal{S}^*_C)$. The sharpness of the result is established using the subsequent function

$$f_0(z) = z\exp\left(\frac{4z+z^2}{3}\right),$$

where $zf_0'(z)/f_0(z) = 1 + (4z+2z^2)/3$ yields $f_0 \in \mathcal{S}^*_\rho$, and substituting $z = R_{\mathcal{S}^*_\rho}(\mathcal{S}^*_C)$ gives $zf_0'(z)/f_0(z) = 1 + \sinh^{-1}(1)$, thereby proving the sharpness.

(iv) Suppose $f \in \mathcal{S}^*_e$, we have $zf'(z)/f(z) \prec e^z$, which yields

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq e^r - 1 \leq \sinh^{-1}(1) \text{ holds in } |z| = r,$$

provided $r \leq \ln(1+\sinh^{-1}(1)) = R_{\mathcal{S}^*_\rho}(\mathcal{S}^*_e)$. Now consider

$$f_0(z) = z\exp\left(\int_0^z \frac{e^t-1}{t}dt\right).$$

Since $zf_0'(z)/f_0(z) = e^z$, where $f_0 \in \mathcal{S}^*_e$, so at $z = R_{\mathcal{S}^*_\rho}(\mathcal{S}^*_e)$, we have

$$zf_0'(z)/f_0(z) = 1 + \sinh^{-1}(1),$$

which shows the sharpness of the result.

(v) Suppose $f \in \Delta^*$ which gives $zf'(z)/f(z) \prec z + \sqrt{1+z^2}$. Then,

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq r + \sqrt{1+r^2} - 1 \leq \sinh^{-1}(1),\ |z| = r,$$

for $r \leq \frac{\sinh^{-1}(1)(2+\sinh^{-1}(1))}{2(1+\sinh^{-1}(1))} = R_{\mathcal{S}^*_\rho}(\Delta^*)$. For sharpness, define f_0 as

$$f_0(z) = z\exp\left(\int_0^z \frac{t+\sqrt{1+t^2}-1}{t}dt\right).$$

Since $zf_0'(z)/f_0(z) = z + \sqrt{1+z^2}$, $f_0 \in \Delta^*$, so at $z = R_{\mathcal{S}^*_\rho}(\Delta^*)$, we have $zf_0'(z)/f_0(z) = 1 + \sinh^{-1}(1)$ which shows the sharpness of the result.

(vi) Suppose $f \in \mathcal{BS}^*(\alpha)$, $\alpha \in [0,1]$, which gives $zf'(z)/f(z) \prec 1 + z/(1-\alpha z^2)$. Then,

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{r}{1-\alpha r^2} \leq \sinh^{-1}(1),\ |z| = r,$$

for $r \leq \frac{-1+\sqrt{1+\alpha\left(2\sinh^{-1}(1)\right)^2}}{2\alpha\sinh^{-1}(1)} = R_{\mathcal{S}^*_\rho}(\mathcal{BS}^*(\alpha))$, $\alpha \in (0,1]$. For $\alpha = 0$, $r \leq \sinh^{-1}(1)$. Next examine the function f_0 defined as

$$f_0(z) = z\left(\frac{1+\sqrt{\alpha}z}{1-\sqrt{\alpha}z}\right)^{1/(2\sqrt{\alpha})}.$$

Since $zf_0'(z)/f_0(z) = 1 + z/(1-\alpha z^2)$, where $f_0 \in (\mathcal{BS}^*(\alpha))$, so at $z = -R_{\mathcal{S}^*_\rho}((\mathcal{BS}^*(\alpha)))$, we have $zf_0'(z)/f_0(z) = 1 - \sinh^{-1}(1)$, which ensures sharpness of the result.

Note that for $\alpha = 0$, $R_{\mathcal{S}^*_\rho}(\mathcal{BS}^*(0)) = \sinh^{-1}(1) \approx 0.881374$ and for $\alpha = 1$, $R_{\mathcal{S}^*_\rho}(\mathcal{BS}^*(1)) = \left(-1+\sqrt{1+(2\sinh^{-1}(1))^2}\right)/(2\sinh^{-1}(1)) \approx 0.58241$. □

Next we present some radius problems for certain classes of functions expressed as ratio of functions:

$$\mathcal{F}_1 := \left\{f \in \mathcal{A}_n : \operatorname{Re}\left(\frac{f(z)}{g(z)}\right) > 0 \text{ and } \operatorname{Re}\left(\frac{g(z)}{z}\right) > 0,\ g \in \mathcal{A}_n\right\},$$

$$\mathcal{F}_2 := \left\{f \in \mathcal{A}_n : \operatorname{Re}\left(\frac{f(z)}{g(z)}\right) > 0 \text{ and } \operatorname{Re}\left(\frac{g(z)}{z}\right) > 1/2,\ g \in \mathcal{A}_n\right\},$$

and

$$\mathcal{F}_3 := \left\{f \in \mathcal{A}_n : \left|\frac{f(z)}{g(z)} - 1\right| < 1 \text{ and } \operatorname{Re}\left(\frac{g(z)}{z}\right) > 0,\ g \in \mathcal{A}_n\right\}.$$

Theorem 4.10. *For functions in the classes $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$, the sharp $\mathcal{S}^*_{\rho,n}$-radii, respectively, are*:

(i) $R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_1) = \left(\frac{\sqrt{4n^2+(\sinh^{-1}(1))^2}-2n}{\sinh^{-1}(1)}\right)^{1/n}$.

(ii) $R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_2) = \left(\frac{\sqrt{9n^2+4\sinh^{-1}(1)(n+\sinh^{-1}(1))}-3n}{2(n+\sinh^{-1}(1))}\right)^{1/n}$.

(iii) $R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_3) = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_2)$.

Proof. (i) Let $f \in \mathcal{F}_1$ and consider the functions $s, d : \mathbb{D} \to \mathbb{C}$, where $s(z) = f(z)/g(z)$ and $d(z) = g(z)/z$. Clearly, $s, d \in \mathcal{P}_n$. As $f(z) = zd(z)s(z)$, applying Lemma 3.2 here gives

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{4nr^n}{1-r^{2n}} \leq \sinh^{-1}(1)$$

such that

$$r \leq \left(\frac{\sqrt{4n^2+(\sinh^{-1}(1))^2}-2n}{\sinh^{-1}(1)}\right)^{1/n} = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_1)$$

holds. Next examine the functions

$$f_0(z) = z\left(\frac{1+z^n}{1-z^n}\right)^2 \text{ and } g_0(z) = z\left(\frac{1+z^n}{1-z^n}\right).$$

Evidently, $\mathrm{Re}(f_0(z)/g_0(z)) > 0$ and $\mathrm{Re}(g_0(z)/z) > 0$, which implies $f_0 \in \mathcal{F}_1$. Further calculation yields at $z = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_1)e^{i\pi/n}$

$$\frac{zf_0'(z)}{f_0(z)} = 1 + \frac{4nz^n}{1-z^{2n}} = 1 - \sinh^{-1}(1),$$

which validates the result is sharp.

(ii) Let $f \in \mathcal{F}_2$ and consider the functions $s, d : \mathbb{D} \to \mathbb{C}$, where $s(z) = f(z)/g(z)$ and $d(z) = g(z)/z$. Clearly, $s \in \mathcal{P}_n(1/2)$ and $d \in \mathcal{P}_n$. As $f(z) = zd(z)s(z)$, applying Lemma 3.2 here gives

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{2nr^n}{1-r^{2n}} + \frac{nr^n}{1-r^n} = \frac{3nr^n + nr^{2n}}{1-r^{2n}} \leq \sinh^{-1}(1),$$

whenever

$$r \leq \left(\frac{\sqrt{9n^2 + 4\sinh^{-1}(1)(n+\sinh^{-1}(1))} - 3n}{2(n+\sinh^{-1}(1))}\right)^{1/n} = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_2).$$

Therefore, $f \in \mathcal{S}^*_{\rho,n}$ holds for $r \leq R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_2)$. Next see that $\mathrm{Re}(g_0(z)/z) > 1/2$ while $\mathrm{Re}(f_0(z)/g_0(z)) > 0$ for the functions

$$f_0(z) = \frac{z(1+z^n)}{(1-z^n)^2} \text{ and } g_0(z) = \frac{z}{1-z^n}.$$

Therefore $f_0 \in \mathcal{F}_2$ which verifies the sharpness for $z = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_2)$ such that

$$\frac{zf_0'(z)}{f_0(z)} - 1 = \frac{3nz^n + nz^{2n}}{1-z^{2n}} = \sinh^{-1}(1).$$

(iii) Let $f \in \mathcal{F}_3$ and consider the functions $s, d : \mathbb{D} \to \mathbb{C}$, where $s(z) = g(z)/f(z)$ and $d(z) = g(z)/z$. Then $d \in \mathcal{P}_n$. We can verify that $|1/s(z) - 1| < 1$ holds whenever $\mathrm{Re}(s(z)) > 1/2$ and therefore $s \in \mathcal{P}_n(1/2)$. As $f(z) = zd(z)/s(z)$, on applying Lemma 3.2, we obtain

$$\left|\frac{zf'(z)}{f(z)} - 1\right| \leq \frac{3nr^n + nr^{2n}}{1-r^{2n}} \leq \sinh^{-1}(1).$$

The rest of the proof is omitted as it is analogous to proof of Theorem 4.10(ii). The sharpness can be verified as follows. Examine the functions

$$f_0(z) = \frac{z(1+z^n)^2}{1-z^n} \text{ and } g_0(z) = \frac{z(1+z^n)}{1-z^n}.$$

Using above definitions of f_0 and g_0, we see that

$$\mathrm{Re}\left(\frac{g_0(z)}{f_0(z)}\right) = \mathrm{Re}\left(\frac{1}{1+z^n}\right) > \frac{1}{2} \text{ and } \mathrm{Re}\left(\frac{g_0(z)}{z}\right) = \mathrm{Re}\left(\frac{1+z^n}{1-z^n}\right) > 0,$$

and therefore, $f_0 \in \mathcal{F}_3$. Now at $z = R_{\mathcal{S}^*_{\rho,n}}(\mathcal{F}_3)e^{i\pi/n}$, we obtain

$$\frac{zf_0'(z)}{f_0(z)} - 1 = \frac{3nz^n - nz^{2n}}{1 - z^{2n}} = -\sinh^{-1}(1),$$

which serves as validation for the sharp result.

This concludes the proof. □

References

[1] I. Aktaş and H. Orhan, *Bounds for radii of convexity of some q-Bessel functions*, Bull. Korean Math. Soc. **57** (2020), no. 2, 355–369. https://doi.org/10.4134/BKMS.b190242

[2] R. M. Ali, N. K. Jain, and V. Ravichandran, *Radii of starlikeness associated with the lemniscate of Bernoulli and the left-half plane*, Appl. Math. Comput. **218** (2012), no. 11, 6557–6565. https://doi.org/10.1016/j.amc.2011.12.033

[3] K. Bano and M. Raza, *Starlike functions associated with cosine functions*, Bull. Iranian Math. Soc. **47** (2021), no. 5, 1513–1532. https://doi.org/10.1007/s41980-020-00456-9

[4] N. E. Cho, S. Kumar, V. Kumar, and V. Ravichandran, *Differential subordination and radius estimates for starlike functions associated with the Booth lemniscate*, Turkish J. Math. **42** (2018), no. 3, 1380–1399.

[5] N. E. Cho, S. Kumar, V. Kumar, V. Ravichandran, and H. M. Srivastava, *Starlike functions related to the Bell numbers*, Symmetry **11** (2019), 219.

[6] N. E. Cho, V. Kumar, S. S. Kumar, and V. Ravichandran, *Radius problems for starlike functions associated with the sine function*, Bull. Iranian Math. Soc. **45** (2019), no. 1, 213–232. https://doi.org/10.1007/s41980-018-0127-5

[7] P. Goel and S. S. Kumar, *Certain class of starlike functions associated with modified sigmoid function*, Bull. Malays. Math. Sci. Soc. **43** (2020), no. 1, 957–991. https://doi.org/10.1007/s40840-019-00784-y

[8] W. Janowski, *Extremal problems for a family of functions with positive real part and for some related families*, Ann. Polon. Math. **23** (1970/71), 159–177. https://doi.org/10.4064/ap-23-2-159-177

[9] S. Kanas and A. Wisniowska, *Conic regions and k-uniform convexity*, J. Comput. Appl. Math. **105** (1999), no. 1-2, 327–336. https://doi.org/10.1016/S0377-0427(99)00018-7

[10] R. Kargar, A. Ebadian, and J. Sokół, *On Booth lemniscate and starlike functions*, Anal. Math. Phys. **9** (2019), no. 1, 143–154. https://doi.org/10.1007/s13324-017-0187-3

[11] W. C. Ma and D. Minda, *A unified treatment of some special classes of univalent functions*, in Proceedings of the Conference on Complex Analysis (Tianjin, 1992), 157–169, Conf. Proc. Lecture Notes Anal., I, Int. Press, Cambridge, MA, 1994.

[12] T. H. MacGregor, *Functions whose derivative has a positive real part*, Trans. Amer. Math. Soc. **104** (1962), 532–537. https://doi.org/10.2307/1993803

[13] R. Mendiratta, S. Nagpal, and V. Ravichandran, *A subclass of starlike functions associated with left-half of the lemniscate of Bernoulli*, Internat. J. Math. **25** (2014), no. 9, 1450090, 17 pp. https://doi.org/10.1142/S0129167X14500906

[14] R. Mendiratta, S. Nagpal, and V. Ravichandran, *On a subclass of strongly starlike functions associated with exponential function*, Bull. Malays. Math. Sci. Soc. **38** (2015), no. 1, 365–386. https://doi.org/10.1007/s40840-014-0026-8

[15] R. K. Raina and J. Sokół, *Some properties related to a certain class of starlike functions*, C. R. Math. Acad. Sci. Paris **353** (2015), no. 11, 973–978. https://doi.org/10.1016/j.crma.2015.09.011

[16] V. Ravichandran, F. Rønning, and T. N. Shanmugam, *Radius of convexity and radius of starlikeness for some classes of analytic functions*, Complex Variables Theory Appl. **33** (1997), no. 1-4, 265–280. https://doi.org/10.1080/17476939708815027
[17] F. Rønning, *Uniformly convex functions and a corresponding class of starlike functions*, Proc. Amer. Math. Soc. **118** (1993), no. 1, 189–196. https://doi.org/10.2307/2160026
[18] G. M. Shah, *On the univalence of some analytic functions*, Pacific J. Math. **43** (1972), 239–250. http://projecteuclid.org/euclid.pjm/1102959657
[19] K. Sharma, N. K. Jain, and V. Ravichandran, *Starlike functions associated with a cardioid*, Afr. Mat. **27** (2016), no. 5-6, 923–939. https://doi.org/10.1007/s13370-015-0387-7
[20] J. Sokół, *Radius problems in the class $\mathcal{SL}^*$*, Appl. Math. Comput. **214** (2009), no. 2, 569–573. https://doi.org/10.1016/j.amc.2009.04.031
[21] J. Sokół and J. Stankiewicz, *Radius of convexity of some subclasses of strongly starlike functions*, Zeszyty Nauk. Politech. Rzeszowskiej Mat. No. 19 (1996), 101–105.
[22] B. A. Uralegaddi, M. D. Ganigi, and S. M. Sarangi, *Univalent functions with positive coefficients*, Tamkang J. Math. **25** (1994), no. 3, 225–230.
[23] L. A. Wani and A. Swaminathan, *Radius problems for functions associated with a nephroid domain*, Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Mat. RACSAM **114** (2020), no. 4, Paper No. 178, 20 pp. https://doi.org/10.1007/s13398-020-00913-4

Kush Arora
Department of Statistics
University of Warwick
Coventry CV4 7AL, UK
Email address: kush.arora1214@gmail.com

S. Sivaprasad Kumar
Department of Applied Mathematics
Delhi Technological University
Delhi-110042, India
Email address: spkumar@dce.ac.in

View publication stats

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

STUDY OF BYPRODUCTS FORMED BY PYROLYSIS OF BIOMASS

Parth Aggarwal[*1], Bhavya Choudhary[*2]

[*1,2]Undergraduate, Department Of Mechanical Engineering, Delhi Technological University, New Delhi, Delhi, India.

ABSTRACT

Pyrolysis is the process of heating organic material, such as biomass, in an inert environment. Since no oxygen is present, combustion doesn't happen, but the chemical compounds, such as cellulose and lignin that make up Biomass thermally decompose into combustible gasses and char. These combustible gasses can be condensed by fractional condensation into a flammable liquid, called pyrolysis oil. The pyrolysis process is used heavily in the chemical industry and mechanical industry, for example, to produce ethylene, many forms of carbon, and other chemicals from petroleum, coal, and even wood, to make coke from coal. It is also used to convert natural gas into non-polluting hydrogen gas and non-polluting solid carbon char, initiating production in industrial volume. Biomass is a sustainable source of fuel and pyrolysis of Biomass results in various solid, liquid, and gaseous products which could be used in a wide array of applications. Biomass pyrolysis has been attracting a lot of attention due to its high efficiency and good environmental performance characteristics. It also provides an opportunity for the processing of agricultural residues, wood waste and municipal solid waste into clean energy.

Keywords: Pyrolysis, Bio-oil, Biochar, Biomass, Fuel.

I. INTRODUCTION

Thermal treatment of waste refers to heat processes used to treat waste materials. These methods for waste management aim to reduce waste volume, convert waste into harmless materials, and utilize the energy hidden within waste as heat, steam, electrical power, or combustible material. Thermal treatment is any waste treatment technology that consists of using high temperatures to process the waste. It is a solution for treating waste in an environmentally and economically friendly way. Thermal treatment reduces the volume and mass of the waste and inerts the hazardous components while at the same time generating thermal or electrical energy and minimizing pollutant emissions to air and water. One of the methods for thermal treatment is pyrolysis. Pyrolysis is a recycling technique converting plastic waste into fuels and monomers by thermal treatment without exposing the waste to oxygen. This paper gives an outlook on pyrolysis and gives a review of pyrolysis in Biomass.

II. PYROLYSIS

Pyrolysis is one of the technologies available to convert biomass to oil, which is mostly referred to as a fuel. It can also be refined to drop in hydrocarbon biofuels, oxygenated fuel, petrochemical replacements, and additives. It is a heating technique of an organic material, such as biomass, in the absence of oxygen.

It is a thermochemical treatment that can be done on any type of organic (carbon-based) product. It can be done on mixers or pure compounds. The process involves heating it at high temperatures in the absence of oxygen through chemical and physical separation into different molecules, mostly monomers.

Further, the decomposition of the material leads to the formation of new molecules. The end product received from this process often has higher characteristics than the original residual; this is an endothermic process that ensures a high energy content in the product received.

Products received after pyrolysis are in solid form. products such as non-condensable gases, biochar, liquid and charcoal (C_nH_m, CO, N, H_2 and CH_4). It is not always true that we receive products in the form of a solid; sometimes we receive the product in the form of a liquid as well, but at a very early stage of the process, the product is released in the form of gas only.

Later on, cooling down, it results in a liquid form. In the middle of the cooling process, the mixture of gas and liquid is called the syngas (a mixture of carbon monoxide (CO) and hydrogen (H_2).

Biomass consists of cellulose, hemicellulose, and lignin, which are complex biopolymers. So, to conclude, we can say the process of pyrolysis defines the composition and yields of the product (pyrolysis oil, char, and syngas).

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

Biomass

All biological organic materials originating from living or recently living creatures are typically referred to as "biomass" but this doesn't apply to fossil fuels like (coal, oil, and natural gas). In terms of energy applications, "biomass" refers to the materials extracted from plants directly, but in a broader aspect, animal-derived materials also fall under its definition. The various forms of biomass are interesting in terms of energy applications, including crops, agricultural forestry residues, wood, industrial, municipal and food waste.

When heat is applied to biomass in a gaseous free/inert environment, it causes various interesting thermochemical reactions. These reactions typically result in the production of a solid residue (char), condensable vapors (bio-oil or tar), and non-condensable gasses. The composition and the yield of the various products are dependent on the type of biomass and the process parameters. In the existing techniques, different terms are applied to the various biomass conversion techniques, differentiated by the heating rate of biomass, the maximum temperature at which the conversion takes place, and the reaction medium.

In Pyrolysis of Biomasses, Woody biomass has been used extensively in the pyrolysis process. The reasons are that it is readily available and relatively homogeneous in comparison to other feedstocks. The utilization of other biomass sources such as agricultural residues, forestry residues, waste, and energy crops could provide cheaper feed sources but also introduce further technical challenges during, before, or after the Pyrolysis Process.

Pyrolysis of Biomass

Biomass consists largely of cellulose, hemicellulose, and lignin, which are complex biopolymers that undergo various transformations as the temperature is increased during the reaction. This can be seen in the thermogravimetric analysis (TGA) results for the different biomass constituents as given in Fig. 4.1.

Initially, the mass of the biomass decreases due to the evaporation of water during the heating stage, i.e., when the biomass temperature is increased to approximately 100 °C. As the temperature increases further to about 160°C, bound H2O is removed from the biomass. The pyrolysis vapours at this stage show a negligible heating value.

At temperatures above 180°C, the complex of cellulose, hemicellulose, and lignin polymers breaks down and liberates a mixture of noncondensable gases and condensable vapors. The condensed tar vapors that contain water, acetic acid, methanol, acetone, phenols, and heavier tars are referred to as pyroligneous acid.

Between 180 °C and 270 °C, the reactions are endothermic, but there have been some instances where the reactions become exothermic at temperatures above 280 °C. It has been suggested that the exothermicity at this temperature range is due to the decomposition of vapours into coke and secondary vapors.

At temperatures above 400 °C, the less volatile components that are still found within the remaining solid char are gradually driven off from the residual, lignin-derived material, thereby yielding a solid product that has a higher fixed carbon content and a lower volatile carbon content.

As the temperature is increased above 600 °C, the primary condensable components that are present in the gas phase undergo cracking and polymerization reactions, resulting in a decrease in the bio-oil yield.

Bio-Oil

Bio-oil (also commonly referred to as "tar") is used to describe the fluid produced from the condensation of the vapours produced during pyrolysis. It is a brownish or blackish coloured liquid depending on the chemical composition of the Bio-Oil, and the heating value of it is approximately equal to that of wood, although it is a high-density liquid of up to 1200 kg/m3. The viscosity of bio-oil can vary from 10 to 1000 cP, depending on the H2O content and the composition of the bio-oil.

The highly polar nature of the bio-oil results in the bio-oil being miscible with water and the phase separation typically occurs at water concentrations of above 30 to 45 wt%. The presence of organic acids in the bio-oil results in a pH of between 2 and 4, and this impacts the downstream processing equipment. The factors affecting the properties of the bio-oil are the biomass feed material (particularly moisture content) and the reaction conditions. Various bio-oil fractions can be obtained by using fractional condensation, which in turn steers the bio-oil properties to some extent.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

Bio-oil is not a steady product, and with time it tends to produce an oil which is more viscous and also allows for the formation of a distinct aqueous phase and a viscous organic phase. This ageing process happens over a period of months at room temperature but can also be accelerated if it is put under higher temperatures where polymerization reactions are promoted.

Bio-oil is the condensed fragments of the cellulose, hemicellulose, and lignin present in the biomass. Bio-oil typically contains a large number of different oxygenated compounds with molecular weights of up to 10,000 g/mol. The large molecules that are present in the bio-oil cannot evaporate at the temperatures that are involved during a normal pyrolysis reaction; therefore, these molecules are entrained from the solid biomass as aerosols. Some molecules present in bio-oil are also the products of secondary degradation reactions, such as oligomers of the primary volatile molecules. Analysis of these large molecules by gas chromatography is not possible as coking occurs when the bio-oil is vaporised after being brought onto the chromatographic column. Additionally, a large fraction will not be evaporated at all at temperatures below 300°C (the usual injection temperature of a GC). High-molecular-weight compounds can also be detected using liquid chromatography, including high-performance liquid chromatography and gel permeation chromatography.

BioChar

Char that is purposefully added to soil as a soil amendment is referred to as "biochar." The qualities of the soil are improved by biochar, which increases fertility and, as a result, increases agricultural yields. According to reports, biochar promotes soil water retention, improves cation exchange capacity (CEC), which is crucial for the retention of nutrients in the soil, and serves as a habitat for microbes.

As shown by multiyear field trials in Brazil, putting crushed, commercially available charcoal to the soil may significantly boost the effectiveness of using mineral fertilizers. Although biochar itself has a poor nutrient value, using it with mineral fertilizer may increase crop yields in a synergistic way. A second advantage of biochar is that it is less biodegradable than biomass and, if stable enough in the soil, might function as a carbon sink for CO_2 abatement efforts.

However, there is disagreement over how biochar affects crop yields and how it behaves as a stable carbon sink, and the scope of these advantages is still being researched. Pyrolysis temperatures of about 750°C optimize the surface areas of biochar, while the choice of input material may have an impact. The selection of the feed material and the occurrence of attrition during pyrolysis and product handling are the main causes of the particle size distribution. The surface chemistry and adsorption of water species in biochar have ramifications for the chemical heterogeneity of biochar that is still being investigated.

The surface chemistry of biochar and the adsorption of aqueous species on its surface are both impacted by its chemical heterogeneity. The retention of nutrients in the soil is significantly influenced by the CEC and anion exchange capacity of biochar. The CEC tends to rise as pyrolysis temperatures rise. The ash-like biochar's reported mineral composition has an impact on how the soil behaves.

III. MODELING AND ANALYSIS

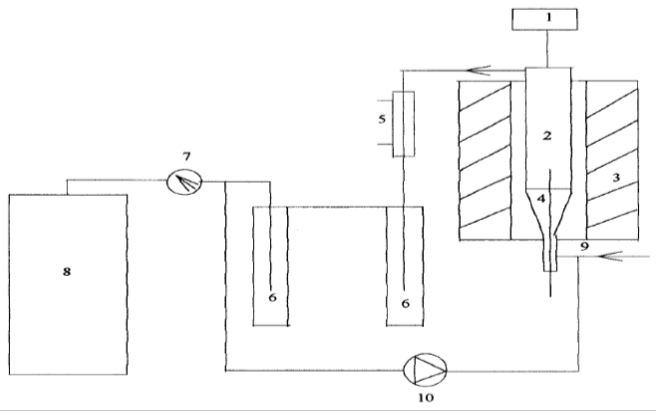


Figure 1: Simplified experimental setting for biomass pyrolysis. (1) Biomass feeder, (2) pyrolysis chamber, (3) electric furnace, (4) packing material, (5) condenser, (6) cold trap, (7) gas volume measure, (8) tight gas tank, (9) sweeping gas stream, and (10) peristaltic pump.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

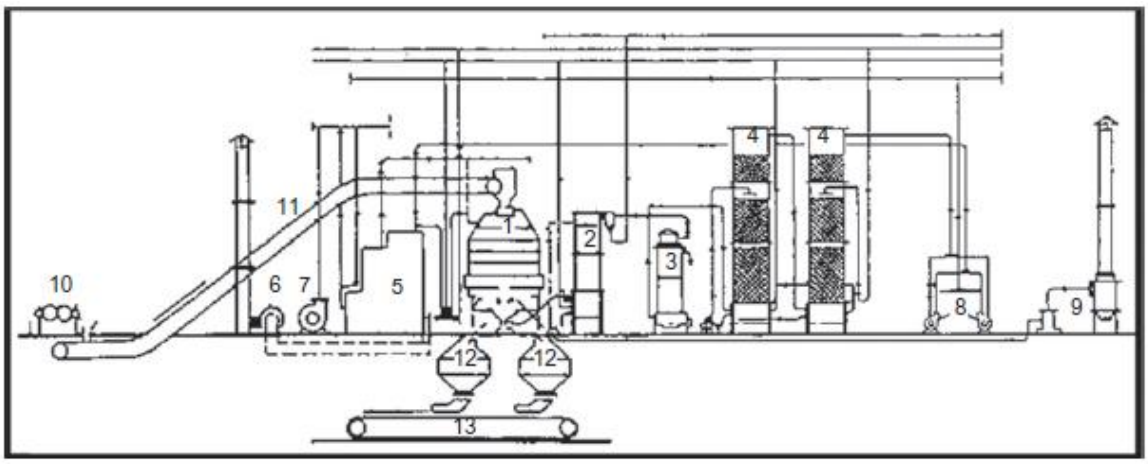


Figure 2: Reichert retort process 1. Large space retort, 2. tar stripper, 3. water cooler, 4. scrubber for residual gas, 5. combustion chamber and heat exchanger, 6. off-gas fan, 7. Combustion air fan, 8. forced gas circulation fan, 9. dust collector, 10. wood preparation, 11. Charging conveyor, 12. charcoal cooler, and 13. charcoal conveyor

IV. TABLES AND DATA

Table 1. Processes for the thermochemical conversion of biomass

Technology	Temperature	Heating rate	Pressure reaction	Reaction time	Medium	Target product
Torrefaction	200–300°C	-	Atmosphere	<2h	Oxygen-free atmosphere	Torrefied biomass
Slow pyrolysis	>300°C	<80°C/min	Atmosphere	Hours - months	Oxygen-free atmosphere	Char
Fast pyrolysis	>300°C	Up to 1000°C/min	Atmospheric or vacuum	< 2s	Oxygen-free atmosphere	Bio-oil
Flash carbonization	<800°C	-	1 MPa	<30 min	Oxygen-free atmosphere	Char
Gasification	600 -1800°C	-	0.1 – 8 MPa	-	Steam and oxygen or air	Gas
Hydrothermal carbonization	180 – 220°C	-	Water saturation pressure	Hours	Water	Char
Wet torrefaction	200 – 260°C	-	Water saturation pressure	5 min	Water	Char
Hydrothermal liquefaction	300 – 350°C	-	15 – 20 MPa	5 – 15 min	Water	Bio-oil

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

Supercritical water gasification	>650°C	-	>22MPa	-	Supercritical water	Gas
Plasma pyrolysis	>3000°C	-	-	-	Ionized gas	Gas

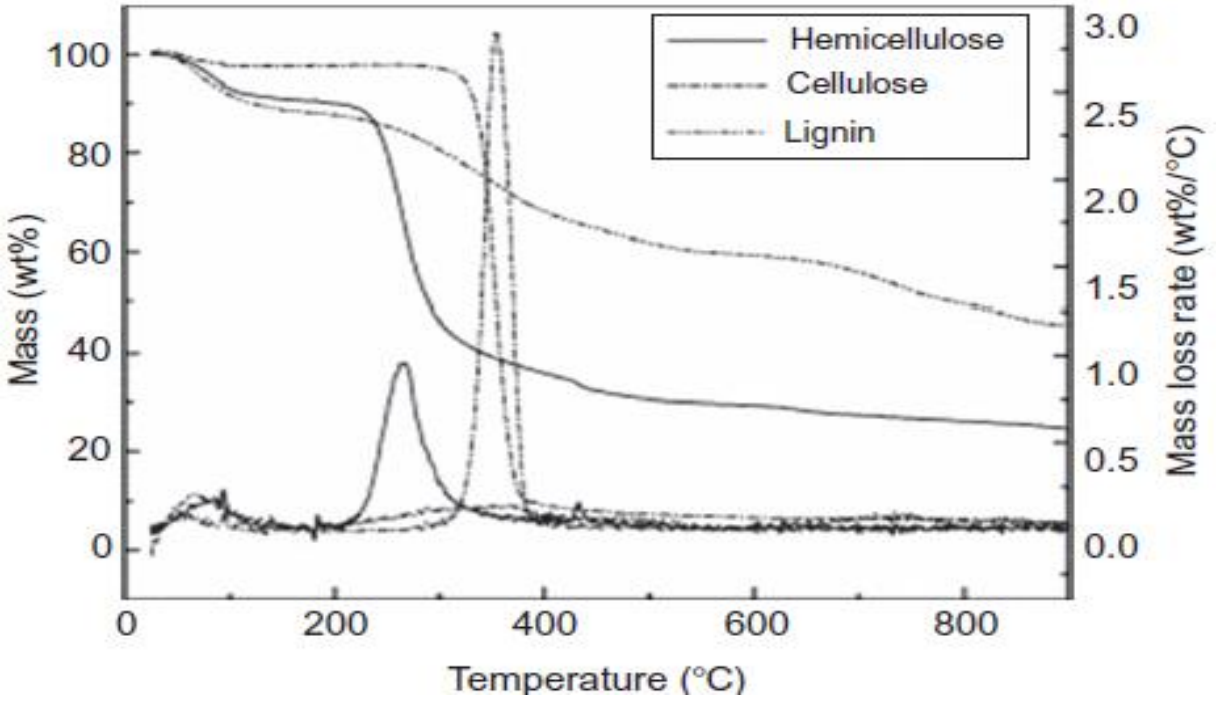


Figure 3: TGA curves for cellulose, hemicellulose, and lignin

V. CONCLUSION

Biomass is a sustainable source of fuel and pyrolysis of Biomass results in various solid, liquid, and gaseous products which could be used in a wide array of applications.

Biomass pyrolysis has been attracting a lot of attention due to its high efficiency and good environmental performance characteristics. It also provides an opportunity for the processing of agricultural residues, wood waste and municipal solid waste into clean energy. In addition, biochar sequestration could make a big difference in the fossil fuel emissions worldwide and act as a major player in the global carbon market with its robust, clean and simple production technology.

Bio-oil can be obtained from the process of fast pyrolysis and its yield can be maximized by conversion at temperatures close to 500°C. The ease of handling liquids and the increased volumetric energy density of bio-oil with respect to biomass are some of the reasons that bio-oil could be perceived as a more useful energy carrier than biomass. The bio-oil may be utilized as a fuel or be processed further to produce chemicals where it is used as a food flavoring.

Biochars can be obtained by low-temperature pyrolysis of biomass wastes. The obtained biochars have a significant increase in carbon content (up to c.a. 70% of C). The improved properties of the char, produced by slow pyrolysis, include increased energy density, hydrophobicity, grindability, and reduced biodegradability when compared to the initial biomass feed, thereby allowing for cheaper transport, handling and processing of the char. Other applications of char include use as a soil amendment.

ACKNOWLEDGEMENTS

We would like to express our sincere and deep gratitude to Dr. Janardan Prasad Kesari, our research supervisor, for their patient guidance, enthusiastic encouragement and useful critiques towards this research work's completion.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:08/August-2022 Impact Factor- 6.752 www.irjmets.com

VI. REFERENCES

[1] Yang H, Yan R, Chen H, Lee DH, Zheng C: Characteristics of hemicellulose, cellulose and lignin pyrolysis, Fuel 86(12–13):1781–1788, 2007.

[2] Ayhan Demirbas & Gönenç Arin (2002) An Overview of Biomass Pyrolysis, Energy Sources, 24:5, 471-482.

[3] Venderbosch R, Prins W: Fast pyrolysis technology development, Biofuels Bioproducts Biorefining 4(2):178–208, 2010.

[4] Food & Agriculture Organization of the United Nations: Industrial charcoal making, Paper 63, 1985.

[5] Bridgwater AV: Review of fast pyrolysis of biomass and product upgrading, Biomass Bioenergy 38:68–94, 2012.

Indian Journal of Engineering & Materials Sciences
Vol. 29, August 2022, pp. 437-444
DOI: 10.56042/ijems.v29i4.48695

Synthesis of diamagnetic ZnO nano-crystallites via sol-gel method and their photocatalytic activity

Deepak Kumar, & Mohan Singh Mehata*

Laser-Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110 042, India

Received: 9 April, 2021; Accepted: 30 July 2021

The diamagnetic zinc oxide (ZnO) nano-crystallites were synthesized via the sol-gel method. The X-ray diffraction (XRD) and field emission scanning electron microscope (FESEM) were used for structural characterization. The XRD pattern demonstrates that the ZnO nanoparticles (NPs) have a wurtzite structure with an average crystallite size of ~20 nm. TheFESEM and EDAX (energy- dispersive x-ray analysis) confirm the particles are in granular, spherical, and rod-like mixed-phase. The absorption spectra show a red shift and observed bandgap energy of about 3.17 eV, which might attribute because of the oxygen vacancies. Further, photoluminescence (PL) exhibits dual peaks at wavelengths 430 and 550 nm, which maybe due to defects induced at the surface of the NPs. The vibrating sample magnetometer (VSM) study demonstrates the diamagnetic behavior of zinc oxide nanoparticles at 300 K. The photocatalytic effectiveness of the ZnO NPs was examined by decomposing a congo red (CR) dye as an organic water pollutant under UV irradiation. The leading absorption peak of CR dye decreases with the UV irradiation time. The complete decolorization of the CR dye was observed after 150 min with the reaction rate constant of 0.018 min^{-1}.

Keywords: Zinc oxide, Photoluminescence, Synthesis, Photocatalytic activity, Diamagnetic behavior, Scavenged

1 Introduction

The material zinc oxide (ZnO) has played a vital role in various sectors of the ever-growing research field. The ZnO is an (II-IV) semiconductor material, possessing wide bandgap energy of 3.3 eV at ambient temperature and pressure with large excitons (60 meV)[1]. It is found in two common polytypes, i.e., cubic zinc and blende hexagonal Wurtzite. The Wurtzite phase having tetrahedral coordination, which exhibits sp^3 covalent bonding, is the most stable phase underfavorable conditions. It has outstanding multifunctional material properties, such as high chemical and photostability, a wide range of radiation absorption, wide bandgap, and cheap and non-toxic. In past decades, researchers paid serious attention to the ZnO for its magnificent properties, which could be employed in various applications like in spintronic, optoelectronic, piezoelectric, and photochemical materials[2], superconductors[3], photocatalyst[4], magnetic[5], light-emitting diodes[6] and biological applications[7].

Various techniques were used to process different types of ZnO nanostructures, including sol-gel technique, co-precipitation, chemical vapour deposition, hydrothermal method, *etc.*[8-9] and solved the purposes. The sol-gel approach proved the most favorable because it is very pocket-friendly, facile, and produces a scalable amount of nanoparticles. The properties and the application of ZnO depend primarily on its shape, size, and structure. The spherical nanostructures of ZnO are mainly exploited for their photocatalytic activities[10].

In this era of industrialization, the increasing demand for products like textiles, toiletries, healthcare, food, beverages, and other industries has led to the release of highly toxic organic compounds like dyes untreated into the water bodies. This causes the destruction of natural reserves of water (rivers, lakes, etc.), which has a long-lasting adverse effect on humanity and the aquatic ecosystem. One such commonly used synthetic dye is the CR dye, an anionic azo and water-soluble red dye used in textile and leather industries. The presence of a meager amount of the compound could lead to chronic carcinogenic diseases in human beings. This industrial pollutant can cause severe environmental effluences such as water pollution, destruction of water plants, and extinction of aquatic life because of the colorization of the water bodies by the dye extracts discarded directly to rivers through sewage.

Over the past few decades, a tremendous increase in the efforts has been employed for eco-friendly

* Corresponding author (E-mail: msmehata@gmail.com)

ways to remove the dye's waste from the water through photocatalytic activity[11]. Many different photocatalytic materials have been developed and used to eliminate CR from sewage waste. The photocatalytic activities of Metal Oxide Semiconductors (MOSs) are preferred because of their efficient adsorption and photocatalytic properties[12]. When the MOSs is transformed into nanoparticles, they exhibit unique physicochemical and biological properties because of the enhancement in surface area to volume ratio.

The ZnO had always been the first choice for photocatalytic decomposition of organic contaminants[13]. The pristine ZnO nanomaterials have been produced via the sol-gel method and employed directly to decompose the congo red dye in the current work. The elaborated structural, morphology and optical study have been investigated for the ZnO-based photocatalytic activity.

2 Materials and Methods

All the high purity chemicals were of analytical grades, such as zinc acetate dihydrate ($(CH_3COO)_2Zn).2H_2O$) bought from Sigma-Aldrich with a purity of 99.99%. Sodium hydroxide (NaOH) has been purchased from the CDH and used without further purification. The de-ionized (DI) water (specific resistivity: 18.2 MΩ·cm) was used as the medium.

4.24 gm of zinc acetate dihydrate powders were dispersed in 100 ml DI and then kept on the magnetic stirrer (900 rpm) for 30 minutes. Simultaneously, 4.0 gm of NaOH was mixed with 50 ml of DI and kept on the magnetic stirrer (800 rpm) for 15 min. We obtained the two clear solutions from the above procedure then the NaOH solution was mixed drop wise in the zinc acetate solution under vigorous stirring for 1 h. The precipitate (white color gel) was formed and left to settle. The residue was washed and filtered four times with DI and ethanol (1:1). It was heated in an oven at 60 °C for 120 min to remove wetness from the precipitate. The dried NPs were calcined at 300 °C for 180 min to get the desire ZnO NPs.

X-ray diffraction (XRD) analysis by Bruker 8D Advance having Cu Kα (1.54 Å) source radiation with an accelerating voltage of 0.04 MV and current of 0.02 Amp. The absorption and reflectance spectra were analyzed with a spectrometer (Perkin Elmer, dual-beam Lambda-750, UV/VIS/NIR). The Photoluminescence (PL) spectra analysis with the Spectrofluorometer Horiba USA (Fluorolog-3), equipped with a double grating at emission and excitation monochromators[14]. The structural morphology and elemental compositions of the samples were obtained using the field emission scanning electron microscope (FESEM, ZEISS Microscope MA15/18) images and energy-dispersive X-ray spectrum analysis (Model EDS 51N1000). VSM (Vibrating Sample Magnetometer) was used to examine the magnetic nature of the produced ZnO NPs.

3 Results and Discussion

The X-ray diffraction patterns depict the structural information about the ZnO, which attributes to the diffraction peaks at Bragg angles of (2θ) 31.76°, 34.42°, 36.25°, 47.53°, 56.59°, 62.85°, 66.374°, 67.94°, 69.08°, that correspond to (100), (002), (101), (102), (110), (103), (200), (112), (201) crystal plane, respectively as shown in Fig. 1. The peaks demonstrated ZnO nanoparticle formation with hexagonal Wurtzite crystal structure, following the standard spectrum of JCPDS Card No. 00-036-1451. The diffraction patterns indicate no impurities exist in the standard hexagonal Wurtzite structure[15]. Using Scherrer's formula Eq. (1), the average crystallite size was estimated (~20 nm) the ZnO peaks ((110), (002))[16].

$$D = \frac{0.9\lambda}{\beta \cos\theta} \quad \ldots (1)$$

where, λ is the X-ray source wavelength (CuKα having λ = 0.154 nm) and 'β' is the full width of the angular peak at the half maximum (Radian), θ is the Bragg's diffraction angle. Using Eq. (2), the lattice parameters, a=b and c, of ZnO nanoparticles were calculated to be 3.245 and 5.196 Å[17].

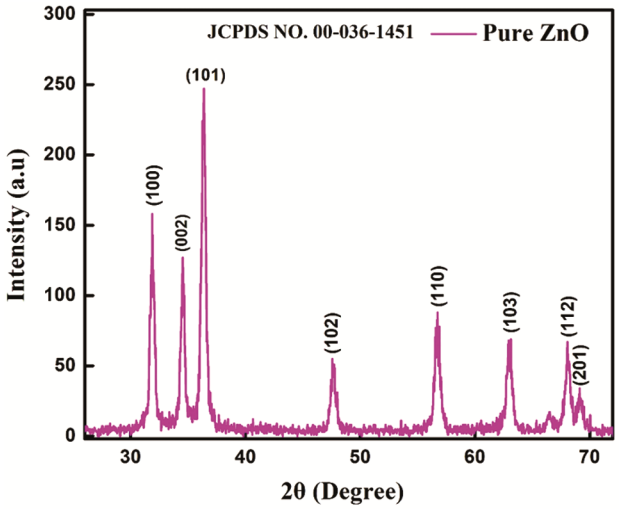


Fig. 1 — X-ray diffraction pattern of ZnO NPs.

$$\frac{1}{d_{hkl}^2} = \frac{4}{3}\left(\frac{h^2+hk+k^2}{a^2}\right) + \frac{l^2}{c^2} \quad \ldots (2)$$

where, the miller indices (h,k,l), inter-planar spacing 'd'. The values of 'a' and 'c' were calculated by (110) and (002) crystal planes corresponding to the peak position of 56.59° and 34.42°. The atomic packing factor (APF) ratio (c/a) and the cell volume of the lattice of the ZnO in the tetrahedral coordination were calculated to be 1.60 and 47.38 (Å)^{3,} respectively[18]. The bond length of all the samples was estimated using Eqs. (3) and (4) came out to be 1.97Å[19].

$$u = \left(\frac{a^2}{3c^2} + 0.25\right) \quad \ldots (3)$$

$$L = \sqrt{\frac{a^2}{3} + \left(\frac{1}{2} - u^2\right)^2 c^2} \quad \ldots (4)$$

The positional parameter 'u' (~0.38) is calculated from Eq. (3). The ionic characteristics and distortion of all the NPs were identified by positional parameters[20]. The crystallite size is inversely proportional to the dislocation density (δ) and how much defect exists in the NPs. The dislocation density (2.5×10^{-3} nm^{-2}) was determined by using Eq. (5)[21],

$$\delta = \frac{1}{D^2} \quad \ldots (5)$$

where, 'D' is the crystallite size obtained from the XRD results. Variation of crystallite size depends on $1/Cos\theta$, while strain with tanθ, and both, in turn, depend on diffraction angle θ. The crystal imperfection and distortion in powder cause strain-induced broadening calculated using Williamson-Hall (W-H) method Eq. (6)[22],

$$\varepsilon = \frac{\beta Cos\,\theta}{4\,Sin\theta} \quad \ldots (6)$$

The microstrain (ε), and dislocation density (δ) of the nanoparticles are more prominent as compared to the crystallite size (D). It may be because of the lattice contractions[23]. The W-H plot of synthesized NPs $\beta cos\,\theta$ (y- axis) vs. $4\,sin\theta$ (x- axis) are presented in the Fig. 2. The degree of distortion 'R' was evaluated by Eq. (7), where a, c are lattice constants.

$$R = \frac{2a(2\div3)^{1/2}}{c} \quad \ldots (7)$$

To verify all the sample's structure is an ideal Wurtzite structure, the degree of distortion value should be unity[24]. The calculated R-values were approximately found to be 1.02

The FESEM images of the nanoparticles are shown in Fig. 3(a) and display the particles are in granular

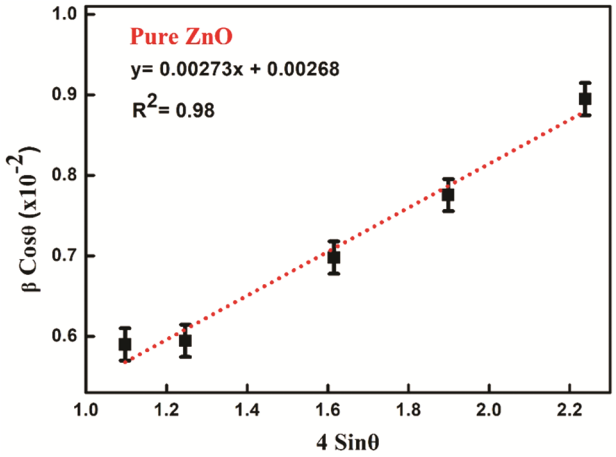
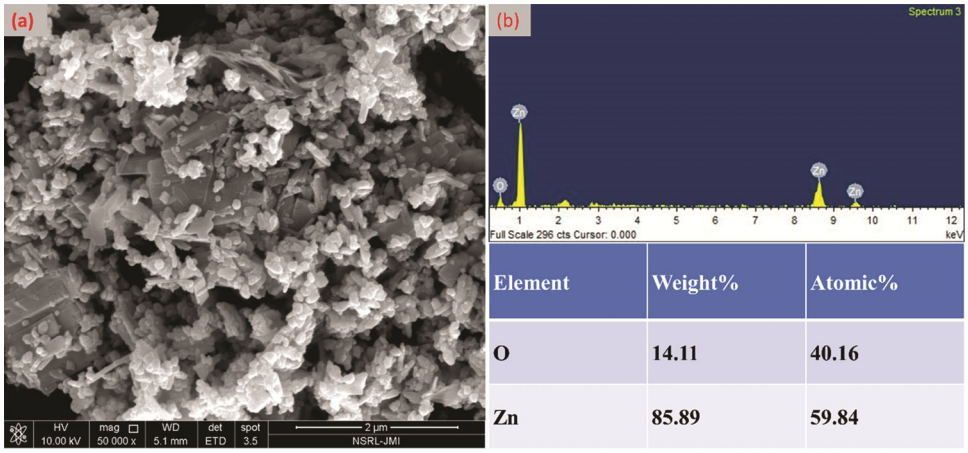


Fig. 2 — The W-H plot between βcosθ *vs.* 4sinθ of synthesized ZnO NPs.



Element	Weight%	Atomic%
O	14.11	40.16
Zn	85.89	59.84

Fig. 3 — ZnO NPs (a) FESEM image, and (b) EDAX spectrum.

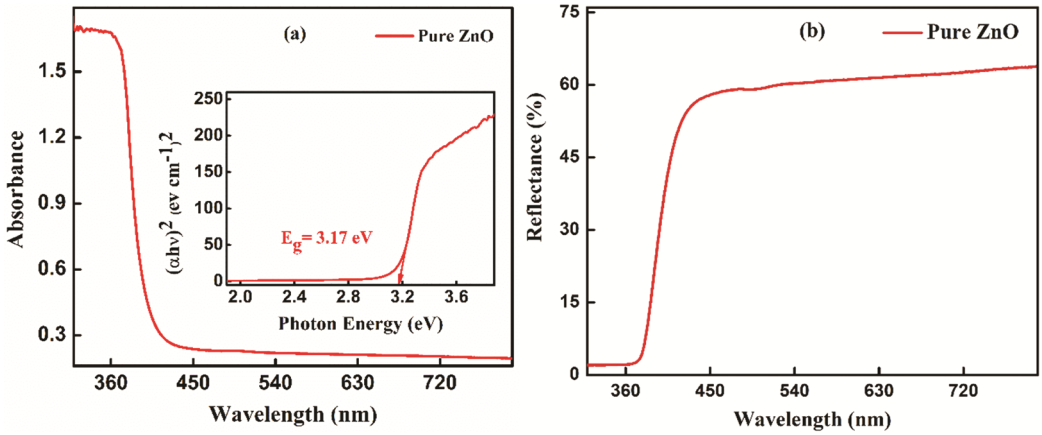


Fig. 4 — (a) The absorption spectrum along with the Tauc plot (inset) corresponds to the direct bandgap, and (b) diffuse reflectance spectrum of the ZnO NPs.

nature, spherical and rod-like mixed-phase (hexagonal and spherical)[25]. The particle size of the ZnO was measured and found to be 54 nm. The elemental compositional analysis was confirmed by the EDAX spectroscopy of the NPs, as shown in Fig. 3(b), which depicts that the Zn and O elements are the leading entities in the sample of ZnO NPs and there were no other fundamental peaks observed in the EDAX spectra.

The absorption and reflectance spectra of ZnO NPs were explored in the wavelength range of 200 to 800 nm at ambient temperature and pressure, as presented in Fig. 4 (a and b). Using the well-known Tauc plot relation Eq. (8), the calculated optical bandgap energy is 3.13 eV[26]. The plot of '$(\alpha hv)^2$ and vs. hv for pure ZnO NPs is shown in Fig. 4 (a).

$$\alpha hv = A\left(hv - E_g\right)^n \quad \ldots (8)$$

E_g represents the bandgap energy of the specimen, α is the absorption coefficient, A is the constant, hv is the incident photon energy, n is the transition coefficient and the value of n depends on the inherent characteristics of semiconductor materials[27]. The values of n = 1/2, 3/2, 2 and 3 represent the direct bandgap, forbidden direct bandgap, indirect bandgap, and forbidden indirect transition gap, respectively.

Nowadays, modern technology is moving towards the design and fabrication section, fabrication of the optoelectronic and integrated circuit (IC) components on the scale of 7, 5, 3 nm and smaller as possible are required. Therefore, it is essential to calculate and compare the bandgap and Refractive Index (RI) of the materials to design and fabricate such sections. Electronic, optical, and optoelectronic applications of semiconductor materials usually depend on their refractive index and bandgap energy. The energy gap measures the photon absorption threshold, whereas the refractive index measures transparency to the incident photons[28]. To determine the refractive index, many empirical relations were proposed[29]. Moss[29] proposed the first relation in 1950 with the known energy gap of semiconductor materials. Five empirical relationships were used to evaluate the refractive index Eq. (9) with the known bandgap (3.17 eV). Suppose the bandgap value of the materials is very small or large. In that case, the relation[29] is not well suited for the calculation of the RI and fits for the range of the energy gap 0.5 eV <Eg< 3.68 eV of the materials[29].

$$n = \left(\frac{95}{\mathrm{Eg}}\right)^{\frac{1}{4}} \quad \ldots (9)$$

Eq. (10)[30-31] validates if the value of energy gap falls in between the range 1.5eV <Eg< 3.50 eV and the theory failed for the values in the region other than this, i.e., if Eg> 6.587 eV and gives negative value[28-30-31].

$$n = 4.084 - 0.63\mathrm{Eg} \quad \ldots (10)$$

The relation Eq. (11) at low and high values of the energy gap[32] provided nearly the same value as obtained from the Moss relation. The expression failed when the energy gap is Eg< 1.4 eV; however, Eq. (10) is reliable in the range of 2.0 eV <Eg< 4.0 eV.

$$n = \sqrt{1 + \left(\frac{13.6}{\mathrm{Eg}+3.47}\right)^2} \quad \ldots (11)$$

The modified Moss relation Eq. (12), in which the constant value of 0.365 was subtracted from the energy gap, yielded a better result than Eq. (9)[33-34]. This relation not valid when bandgap is about 0.365 eV but valid for the range of 1.10 eV <Eg< 6.20 eV.

$$n = \left(\frac{154}{\mathrm{Eg}-0.365}\right)^{\frac{1}{4}} \quad \ldots (12)$$

Kumar and Singh[35] reported in Eq. (13) that when the energy gap is at a lower value, its refractive index lies between the Moss and Reddy relation. While at a higher value of Eg, it lies below the Reddy relation and higher in the Moss relation. The typical reliable rage is 2.00 eV <Eg< 4.00 eV[36-37].

$$n = 3.3668(\mathrm{Eg})^{-0.3228} \quad \ldots (13)$$

Above all, the relation gives a better result when the energy gap lies between the ranges of 2.0 eV to 4.0 eV. Thus, the refractive index values of the ZnO NPs were calculated from the given above five relations Eqs (9 - 13). The refractive index estimated from Moss relation is (2.33) whereas from Kumar & Singh[35] to be (2.32), nearly the same. The highest refractive index of NPs is 2.72 was calculated through the Reddy and Anjayenulu[23] relations, while the lowest refractive index (2.08) of NPs by Ravindra *et al*[31]. The moderate RI was calculated by Harve *et al*[32]. is (2.27).

PL spectra provide information about excited states and their energy and the presence of defects and trap states. The spectral intensity is associated directly with the defect density or structural defects in the semiconductors[38]. The PL spectrum of ZnO NPs has been measured in the wavelength range of 200 to 800 nm at the excitation wavelength of 320 nm at ambient conditions and illustrated in Fig. 5. A broad PL spectrum is observed with a band maximum at 540 nm. The Gaussian bands (G1 and G2) well reproduce the PL spectrum. G1 appeared at 434 nm and G2 at 540 nm. These bands are well-matched with the reported values[39-40]. The G1 band, which is located at around 434 nm, originated from the direct bandgap energy, i.e., the recombination of electrons (e^-) from the conduction band (CB) to the zinc vacancies (V_{zn})

Whereas the relative stronger band observed at 540 nm related to the oxygen vacancies (V_O) and zinc interstitials (Zn_i), or transition from the valence band to the oxygen vacancies (V_O) and zinc interstitials (Zn_i).

The magnetic characteristics of the synthesized ZnO NPs were measured at 300 K via a vibrating sample magnetometer. The M-H curve of the NPs exhibited diamagnetic nature, as shown in Fig. 6. The purely diamagnetic behavior of the ZnO at the low applied magnetic field supports the presence of oxygen vacancies (V_O) and zinc vacancies (V_{zn}), as described by PL spectrum. Besides, the lattice distortion and the size effect can also be responsible for the diamagnetic magnetic behavior of the NPs[41].

The photocatalytic activity (PCA) of ZnO was tested for the photodegradation of an anionic congo red (CR) azo dye under UV irradiation. For the study, the dye concentration was chosen to be 5.0×10^{-5} M. Figure 7 (a) shows the absorption of the CR dye in

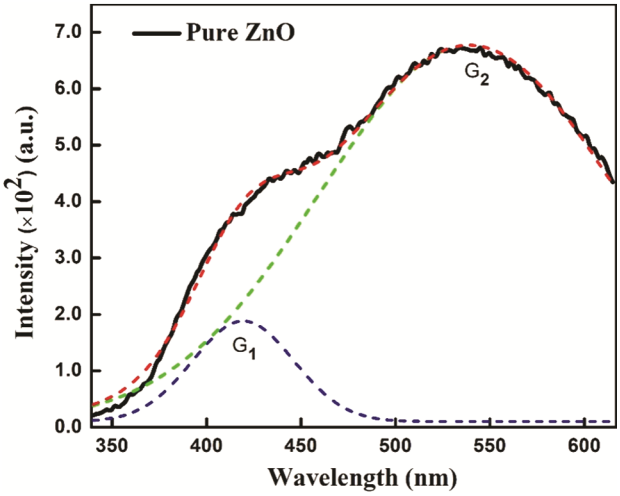


Fig. 5 — The photoluminescence spectrum (solid red line) of ZnO NPs and G_1 and G_2 Gaussian bands (dotted blue and green line) along with the fitted curve (red dotted line). The excitation wavelength (λ_{ex}) was 320 nm.

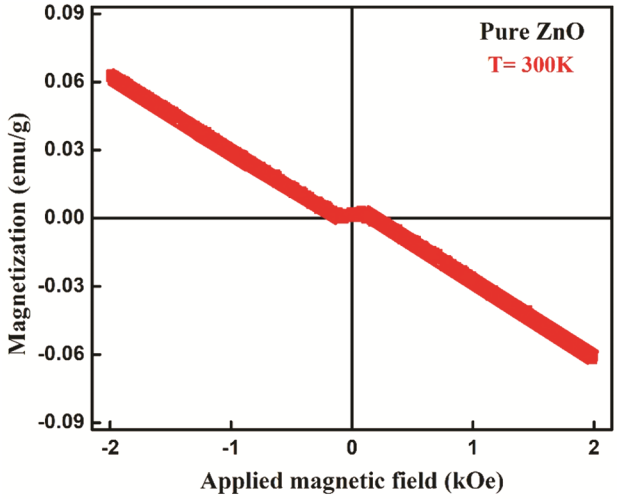


Fig. 6 — M-H curve of ZnO NPs at the 300 K.

water. The maximum absorption corresponding to the lowest energy band is at 498 nm. The blank CR dye solution (without a mixture of ZnO catalyst) was kept under UV irradiation to examine the photolytic reaction. Still, there was no effect, hence no decomposition of dye without catalysts. Then, the typical dye degradation was performed in the presence of ZnO NPs as a catalyst (0.8 gL^{-1}) in an aqueous medium.

In this process, the catalyst (ZnO NPs) was initially added to the dye solution and stirred in the dark for 1 h to reach the adsorption-desorption isotherm. Then 4 ml of the mixed solution was extracted from the sample and centrifuged at 7000 rpm for 2 min. The sample's absorption spectra were recorded before and after the exposure with UV illumination at specific intervals. Before photoirradiation, the CR dye was degraded slowly, and the rate became extremely slow after some intervals, i.e., the change in absorption is nearly minimal (Fig. 7(a)). After that, the specimen was kept under a UV lamp of 8 Watts, λ ~ 254 nm, and 1.2 mW/cm^2. Upon the photoirradiation, the absorption intensity of CR dye decreases; hence the photodegradation of the dye starts. Upon increasing the photoirradiation time, the absorption intensity gradually decreased and became minimal at an irradiation time of 150 min. The reaction rate constant of dye degradation has been calculated using pseudo-first-order kinetic relation $ln\ (A/A_0) = -kt$[42]. Where 'A_0' and 'A' are the dye absorbance at initial and after a specific irradiation time, respectively. The 'k'is the pseudo-first-order rate constant (min^{-1}), and 't' is the photo-irradiation time (min). The reaction rate constant response is 0.018 min^{-1} obtained from the linear fitted curve. The linear fitted curve is plotted between the $ln\ (A/A_0)$ vs. t (irradiation time) as shown in Fig. 7(b). The degradation efficiency (DE) was estimated by using Eq. (14)[43].

$$\text{Degradation efficiency (DE)}\% = \frac{A_0 - A}{A} \times 100 \quad \ldots (14)$$

The photodegradation efficiency of pure ZnO was calculated to be 98 %.

In the process of PCA, the electron-hole pairs are generated by the exposure of adequate UV light energy (hv) on the ZnO nanoparticles. The generated holes and electrons participate in the decomposition of CR dye, as demonstrated in Fig. 8.

$$ZnO + hv(\text{UV irradiation}) \rightarrow ZnO\ (e^- - h^+) \quad \ldots (15)$$

To produce hydroxyl (·OH), the holes (h^+) are formed when the valence band electrons of the catalyst react with electron donors, i.e., H_2O and hydroxyl ions (OH^-)[44]. In the conduction band, the electrons react with the O_2 (oxygen molecule) on the ZnO surface to produce reactive super radical anions ($\cdot O_2^-$)[45]. In the formation of hydroperoxyl radical ($HO_2\cdot$) and hydrogen peroxide (H_2O_2), the reaction will continue, and finally, the hydroxyl radicals are formed. Thus, the hole (h^+) and hydroperoxyl radicals are highly reactive intermediates attacking dye repeatedly and eventually decomposing the dye or organic pollutants[46-47].

$$\text{Dye} + (HO_2\cdot), (\cdot O_2^-), (OH^\cdot) \rightarrow \text{intermediates} \rightarrow \text{final products} \quad \ldots (16)$$

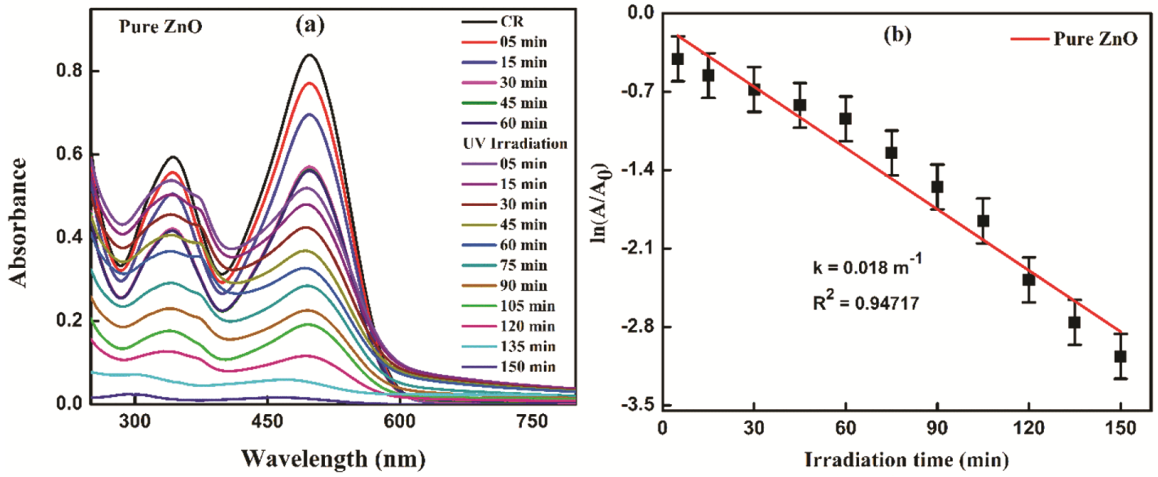


Fig. 7 — (a) The absorption spectra of CR dye in water in the presence of pure ZnO under UV irradiation at different time intervals , and (b) plot of $\ln(A/A_0)$ vs. t (irradiation time).

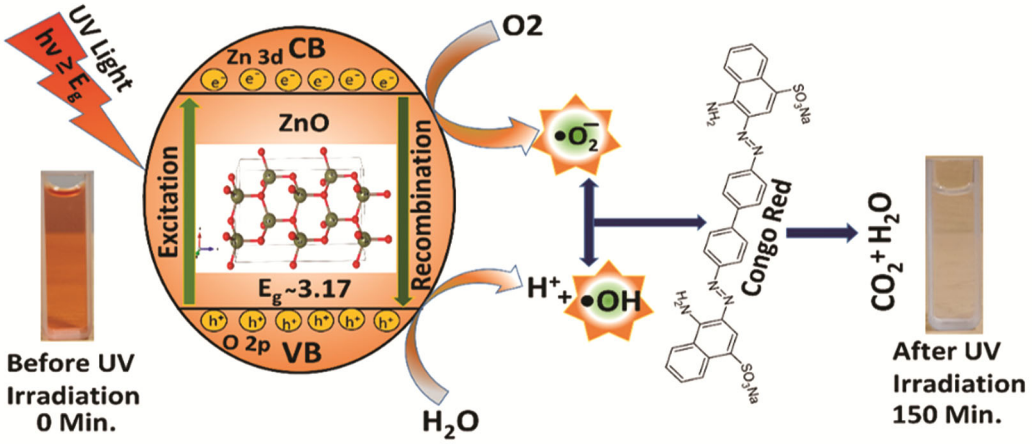


Fig. 8 — Schematic representation of the photocatalytic process and dye degradation in the presence of ZnO nanoparticles.

Nevertheless, trapping e^- in the CB to be scavenged through the e^- acceptors suppresses its recombination with the trapped hole.

4 Conclusion

The pure ZnO NPs were produced via the facile sol-gel method. The structural and elemental investigations of the NPs were carried out with XRD, FESEM with EDAX, UV-vis, PL and VSM. The observed crystallite size of the synthesized NP was ~20 nm. FESEM and EDAX images depict the particles in granular, spherical, and rod-like mixed-phase, and Zn and O elements exist in the ZnO NPs. The bandgap of NPs is ~3.17 eV obtained from the absorption spectrum. The PL spectrum demonstrates that the defects are present in the ZnO, which may be oxygen or zinc vacancy due to which diamagnetic behavior appears in the ZnO NPs. The CR dye degraded under UV irradiation. Within 150 min, the specific amount of the dye degraded entirely with 98% efficiency.

Acknowledgment

DK is thankful to all the lab members for their help, especially Dr. Prateek Sharma.

References

1 Kumar Y, Sahai A, Olive-Méndez S F, Goswami N, & Agarwal V, *Ceram Int,* 42 (2016) 5184.
2 Kariper I A, *Opt Mater,* 44 (2015) 78.
3 Li Y J, Wang C Y, Lu MY, Li K M, & Chen L J, *Cryst Growth Des,* 8 (2008) 2598.
4 Ghosh S S, Choubey C, & Sil A, *Superlattices Microstruct,* 125 (2019) 271.
5 Mesaros A, Ghitulica C D, Popa M, Mereu R, Popa A, Petrisor T, Gabor M, Cadis A I, & Vasile B S, *Ceram Int,* 40 (2014) 2835.
6 Alsharafi R, Zhu Y, Li F, Xu Z, Hu H, & Guo T, *Org Electron,* 75 (2019) 105411.
7 Lima M K, Fernandes D M, Silva M F, Baesso M L, Neto A M, de Morais G R, Nakamura C V, de Oliveira Caleare A, Hechenleitner A A W, & Pineda E A G, *J Sol-Gel Sci Technol,* 72 (2014) 301.
8 Devi P G, & Velu A S, *J Theor Appl Phys,* 10 (2016) 233.
9 Sutanto H, Wibowo S, Hadiyanto, Arifin M, & Hidayanto E, *Mater Res Express,* 4 (2017) 076409.
10 Güy N, Çakar S, & Özacar M, *J Colloid Interface Sci,* 466 (2016) 128.
11 Kumar D, Singh M K, & Mehata M S, *Mater Res Bull,* 150 (2022) 111795.
12 Jamal N, Radhakrishnan A, Raghavan R, & Bhaskaran B, *Main Gr Met Chem,* 43 (2020) 84.
13 Huang Y T, & Shih M C, *Int J Sci Res Publ,* 6 (2016) 549.
14 Mehata M S, & Ratnesh R K, *Dalt Trans,* 48 (2019) 7619.
15 Ji Huang, Cai C, Zhou S, & Liu W, *J Mater Sci Mater Electron,* 29 (2018) 12917.
16 Singh M K, & Mehata M S, *Optik*, 193 (2019) 163011.
17 Hassan F, Miran M, Simol H, Susan M H, & Mollah M, *Bangladesh J Sci Ind Res,* 50 (2015) 21.
18 Islah-u-din, Salam W T, Gillani S S A, Tahir M B, Ikram M, & Ali S, *Optik,* 203 (2020) 163966.
19 He R, Tang B, Ton-That C, Phillips M, & Tsuzuki T, *J Nanoparticle Res,* 15 (2013) 1.
20 Gupta G, Verma S, Nagarajan R, & Rath S, *Phys B & Condens Matter,* 604 (2021) 412735.
21 Srinivasulu T, Saritha K, & Reddy K T R, *Mod Electron Mater,* 3 (2017) 76.
22 Khorsand Zak A, Majid W H Abd, Abrishami M E, & Yousefi R, *Solid & State Sci,* 13 (2011) 251.
23 Tiwari A, & Sahay P P, *Opt Mater*, 110 (2020) 110395.
24 Ran A K, Kumar Y, Rajput P, Jha S N, Bhattacharyya D, & Shirage P M, A*CS Appl Mater Interfa,* 9 (2017) 7691.
25 Thirupathi C, Nithiyanantham S, Sentilkumar M, Arivudainambi A, Mahalakshmi S, & Natarajan B, *Adv Sci Eng Med,* 12 (2020) 524.
26 Singh M K, & Mehata M S, *Ceram Int,* 47 (2021) 32534.
27 Fang M, Tang C M, & Liu Z W, *J Electron Mater,* 47 (2018) 1390.
28 Tripathy S K, *Opt Mater,* 46 (2015) 240.
29 Moss T S, *Phys Status Solidi,* 131 (1985) 415.

30 Gupta V P, & Ravindra N M, *Phys Status Solidi,* 100 (1980) 715.
31 Ravindra N M, Auluck S, & Srivastava V K, *phys stat sol,* 93 (1979) K155-K160.
32 Hervé P, & Vandamme L K J, *Infrared Phys Technol,* 35 (1994) 609.
33 Reddy R R, & Anjaneyulu S, *Phys Status Solidi,* 174 (1992) K91.
34 Reddy R R, & Ahammed Y N, *Infrared Phys Technol,* 36 (1995) 825.
35 Kumar V, & Singh J K, *Indian J Pure Appl Phys,* 48 (2010) 571.
36 Kaphle A, & Hari P, *J Appl Phys,* 122 (2017) 165304.
37 Senol S D, Ozugurlu E, & Arda L, *J Alloys Compd,* 822 (2020) 153514.
38 Salameh B, Alsmadi A M, & Shatnawi M, *J Alloys Compd,* 835 (2020) 155287.
39 Bhardwaj R, Chae K H, & Goyal N, *Vacuum,* 178 (2020) 109446.
40 Dasari M P, Godavarti U, & Mote V, *Appl Ceram,* 12 (2018) 100.
41 Chanda A, Gupta S, Vasundhara M, Joshi S R, Mutta G R, & Singh J, *RSC Adv,* 7 (2017) 50527.
42 Singh M K, & Mehata M S, *Opt Mater,* 109 (2020) 110309.
43 Debnath P, & Mondal N K, *Environ Nanotec & Monit Manag,* 14 (2020) 100320.
44 Lam S M, Sin J C, Abdullah A Z, Mohamed A R, & Desalin, *Water Treat,* 41 (2012) 131.
45 Sadollahkhani A, Hussain Ibupoto Z, Elhag S, Nur O, & Willander M, *Ceram Int,* 40 (2014) 11311.
46 Sharma P, Singh M K, & Mehata M S, J *Mol Struct,* 1249 (2022) 131651.
47 Adam R E, Pozina G, Willander M, & Nur O, *Photo Nanostr & Fundam Appl,* 32 (2018) 11.

The compressive study of energy security prospects in India through Solar Power

Om Prakash[1], Asim Ahmad[2], Anil Kumar[3], Rajeshwari Chatterjee[4], Shubham Sharma[5,6*], Reza Alayi[7*], Hossein Monfared[8]

[1]Department of Mechanical Engineering, Birla Institute of Technology Mesra, Ranchi, India.

[2]Faculty of Engineering and Applied Sciences, Usha Martin University, Ranchi, India.

[3]Department of Mechanical Engineering, Delhi Technological University, Delhi, India.

[4]Department of HMCT, Birla Institute of Technology Mesra, Ranchi, India.

[5]Department of Mechanical Engineering, University Centre for Research & Development, Chandigarh University, Mohali, 140413, India;

[6]Deptt. of Mechanical Engg., IK Gujral Punjab Technical University, Main Campus-Kapurthala 144603, India;

[7]Department of Mechanics, Germi Branch, Islamic Azad University, Germi, Iran.

[8]Department of Mathematics, Germi Branch, Islamic Azad University, Germi, Iran.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

© The Author(s) 2022. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution License https://creativecommons.org/licenses/by/4.0/, which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Energy security depends upon the supply and demand of electricity. Power supply volatility is a challenge in India power sector. India is mainly dependent on coal, which is not a viable long-term option. Hence the future of energy supply in India is renewable energy, and solar energy is the most prominent and reliable source of renewable energy. The solar power plant is classified as a low, medium, and high temperature based solar power plant. The low- temperature solar power plant, the collectors, are unglazed plates. In contract, medium-temperature collectors are flat plates, and high-temperature collectors concentrate sunlight using mirrors and lenses, achieving temperature and pressure up to 3000^{o}C and 20 bar. This article discusses the existing total energy scenario, current solar energy developments, supportive solar energy policies, and solar energy prospects. The tactics, availability, perspectives, potential, and successes of solar energy in India are discussed in this study. The authors compared the barriers to solar energy acceptance and examined the topic of social dissimilarity from a broad theoretical perspective. The study offers policymakers at the federal and state levels, as well as investors, a perspective on issues that may lead to investments in this area. The study states that solar energy could meet above 50 % of electricity sector demand in India in 2040. This study determined that solar energy incidence in India is around 5000 trillion KWh (kilowatt-hours) each year. The solar energy accessible in a single year surpasses the energy output from the petroleum derivative. The average energy from the solar power plant is 0.30 kWh per m^2 equal to the 1400–1800 peak rated capacity. India has many solar power facilities, making it one of the top producers of renewable energy power.

Keywords: Fossil Fuels; Renewable Energy; Solar Energy; Energy security; Electricity

Abbreviation

AD = Accelerated Depreciation

C-WET = Centre for Wind Energy Technology

CERC = Central Electricity Regulatory Commission

CSP = Concentrated Solar Power

NCEF = National Clean Energy Fund

NSM = National Solar Mission

PV = Photovoltaic

PPA = Power Purchase Agreement

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

CC = concentrating collectors

D = Distribution

FIT = Feed-In Tariff

GBI = Generation Based Incentive

GDP = Gross Domestic Product

HTF = Heat transfer fluid

ISA = International Solar Alliance

IREDA = Indian Renewable Energy Agency

JNNSM = Jawaharlal Nehru National Solar Mission

MNRE = Ministry of new and renewable energy

MOP = Ministry of Power

RECL = Rural Electrification Corporation Limited

RPO = Renewable Purchase Obligation

RECs = Renewable Energy Certificates

RPSSGP = Rooftop PV and Small Solar Power Generation Program

RRF = Renewable Regulatory Fund

SHP = Small Hydropower

SPV = Solar Rooftop Photovoltaic

SERCs = State Electricity Regulatory Commissions

T = Transmission

UT = Union territories

1. Introduction

The power shortage is the difference between the demand and supply of electricity. In India, there is a huge gap between energy demand and supply. Most of the rural and agricultural sectors of India is either using free electricity or at very high subsidized rates and thus resulting in a huge gap in electricity supply and the power cost. The power consumption in rural areas is also not metered. It makes difficult in knowing the average total power requirement in a month [1]. The estimation of power consumption is also not accurate and transparent. India is majorly dependent upon coal for thermal power generation which is approximately 67 % of total power generation [2]. The major issue lies in distribution, technical losses due to theft, distorted tariff structure. [3]. The investment in renewable sources like solar and wind is likely to pay off the demand and supply gap [4]. The cost of power generation from solar energy is low due to

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

abundant availability of solar energy [5]. India's focus on solar energy would address both energy security and environmental concerns. Most parts of India have 300 days of sunshine; therefore, solar energy could help to achieve energy security in a prominent way [6].

The central government of India has launched Jawaharlal Nehru National Solar Mission (JNNSM) scheme to promote solar energy [7]. The JNNSM revised to achieve a target of 100 GW of power generation of solar PV(Photovoltaic) by 2022 [7]. JNNSM's goal is to make India a global centre for solar energy by promoting its implementation and policy throughout the country. Differing key policies and targets support each phase. Phase 1 was from 2010 to 2013, the target of 1000 MW for grid-connected PV including rooftop and 200 MW off-grid connection [8]. Phase 2 was from 2014 to 2017. The cumulative target for grid-connected solar PV is 4000 – 10000 MW and 1000 off-grid connections [9]. According to Ultra Mega Solar Power Projects, the target is 40 GW solar PV by year 2022 [10]. Phase 3 is from 2017 to 2022. The target for grid-connected solar PV, including rooftop, is 100000 MW and for off-grid solar PV applications is 2000 MW [8-10]. The International Solar Alliance (ISA) an important global initiative to boost solar capacity globally. India would be self-sufficient in a solar power way 25 to 30 years from now [11].

The obstacles to achieve the power security through solar power plant are grid connection, financing cost, insufficient subsidy, high interest rate and lack of grid capacity [12]. Solar Intensity is the biggest hurdles for adoption of solar power since panel’s output depends on the amount of solar intensity received [13]. Thus, solar is a much better power source in southwest parts of the country. The payback period is higher for regions without intense solar coverage. The efficiency of a solar panel cannot convert entire power into useful power [14-15]. The widely available solar panels have efficiency ratings of less than 26 % [15]. The tracking panel follows sun but finally reach peak power time. India has enough policies and regulations to overcome this setback through proper implementation.

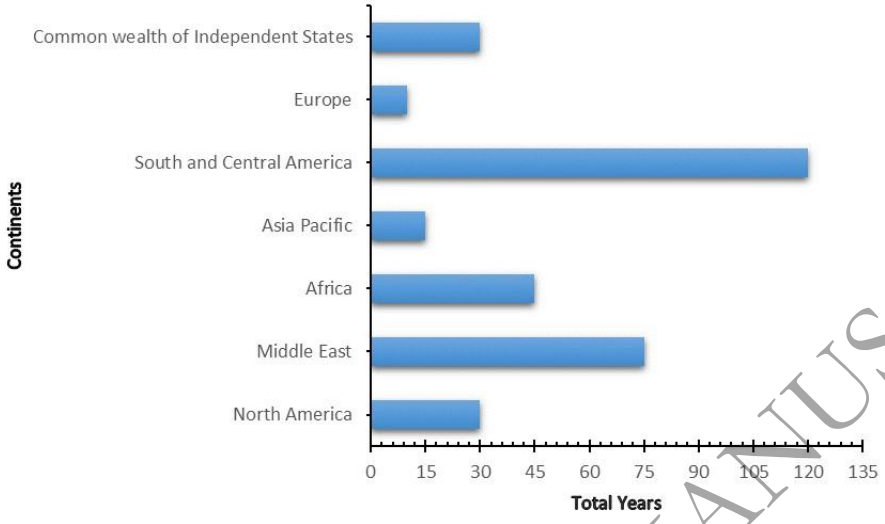
In 2017, it was reported that the reserve/production ratio for fossil assets in Europe, America, and Asia (Pacific) was 10.4, 30.8 and 16.7 in years respectively, depicted in Fig. 1 [16].

The solar energy has a beneficial advantage over wind energy, as it available in enormous amount, large surface of the earth and is easily predictable. For example, wind speed can change from 45 to 190 km/h in 4s, which is quite a notable change thus it is difficult for engineers to design suitable wind turbine [17].

ORIGINAL UNEDITED MANUSCRIPT

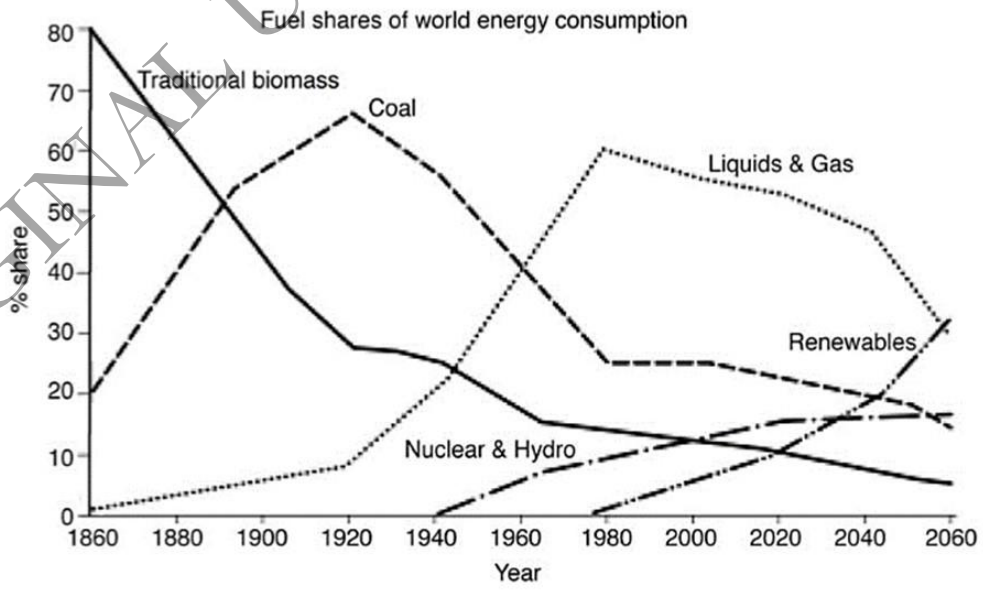
Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Fig. 1. Reserves to production (R/P) ratio estimates [16].



India's population is expected to double in 2051 (half of the 21st century) [17]. Enormous population and growth activities would lead to an enormous amount of energy consumption. Hook et. al (2012) represented the futuristic growth of renewable energy [18]. This prediction is highlighted in Fig. 2 indicating the traditional biomass and coal production share will tend to decrease and would be replaced by renewable energy, including solar hydro and nuclear in 2060 [19].

Fig. 2. Futuristic and historical trends of energy sources [19].



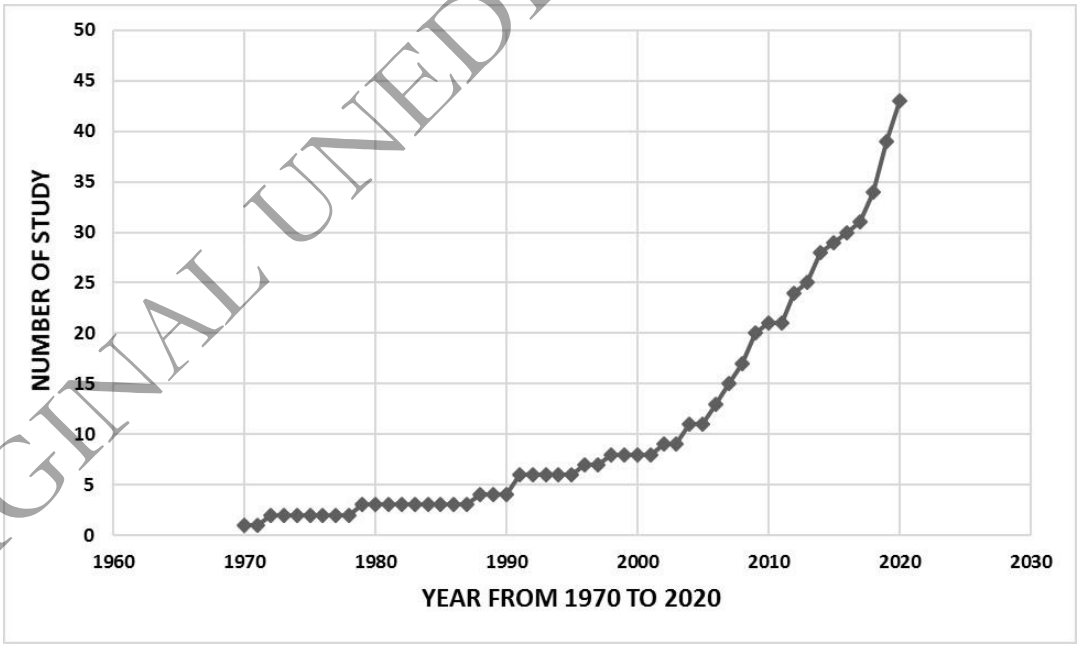
ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

The study on solar energy is required to fulfill the power of main energy consuming Indian cities. India aims to achieve 175 GW of renewable energy, consisting of 100 GW solar energy, 60 wind power, 10 GW bio-power and 5 GW hydropower plants by 2022 [20-23]. The recent investment plan is to achieve more than 270 GW power generation [24]. The previous study shows that in 2047 solar potential would be 750 GW [25].

The many studies and research done on technical feasibility solar power. The more focus laid on renewable energy while missing out the major prospect of solar power. The study is conducted every year but still infective to give a holistic approach and view of solar energy security in India. The estimated study every year is summarized in Fig. 3 indicating the solar energy increased linearly. The future prospect and energy forecasting is difficult to analyze due to changing environmental problem, emission, and regulation. The previous research missed importance of solar power generation through low temperature and high temperature concentrating collector since solar panel efficiency is limited to only 25 % [26]. Moreover, researchers have missed the importance of solar policy to achieve energy security.

Fig. 3. Number of solar energy study in India.



The present study aims to evaluate the current energy status in India and addresses the technical fitness of solar energy to plan for India. The stress is laid upon solar energy.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

The novelty of this paper describe the strategies, availability, perspectives, potential and achievements of solar energy in India. The authors have made an attempt for the first time to compare the barriers towards the acceptance of solar energy. The Author also analyzed on the social dissimilarity issue from a broad theoretical perspective.

The study gives a view to the policymakers at the central and state government and investors, on issues that can potentially lead to investments in this sector. This paper presents the current total energy scenario, solar energy developments in the present stage, supportive solar energy policy and prospects of solar energy.

2. Objectives

The solar energy policy in India has been regularly improved to attain energy security. Due to this, the generation through the solar power plant is being increased rapidly. However, there is still a long way to take full advantage of the solar power in the Indian scenario.

Thus, it is imperative to understand the techno-economical aspect of solar power projects. The current obstructions to the solar power deployment emerging because of past energy policy needed to be recognized and solved for quick deployment. The goal of the current work is to present the comprehensive status of the solar power plant, and the techno-economic barriers while building up a sustainable solar power plant so India can attain energy security. The objectives of the paper categorized are as follows:

I. To present the renewable energy assets, their future contribution, growth, and demand
II. To provide an overview of the current energy scenario, energy demand growth and scarcity in India
III. To provide a scenario of current solar power utilization and future prospect in India
IV. Finally, to present the current renewable energy contribution and government policies

3. Energy Situation in India

India is a fast-growing developing country with an enormous population of about 1.39 billion until 2020 [27]. Most of the people (almost half of the population) are middle-class people. It's per capita growth of energy is about 3629 KW-h [28]. However, about 580 million people lack electricity in India.The addition of energy-producing units is lower in India as compared to other developed countries. This leads to power shortages and an increased cost of electricity per unit [29]. Hence government has to spend huge amounts of wealth to give subsidies/loans in the electricity for the needy middle-class person [30]. Thus, Government of India has formed "Rural

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Electrification Corporation Limited (RECL)", and this is to provide noteworthy economic support for the needy citizen. The board have sanctioned loans about INR $83,871\times10^7$ in year 2016-17; INR $1,07,534\times10^7$ in year 2017-18; INR $1,15,957\times10^7$ in year 2018-19; INR $1,10,908\times10^7$ in year 2019-20 [31-33]. Loan is also sanctioned in area of renewable based power plant installation, generation, distribution, and transmission [34]. Annual gross electricity generation is represented in Table 1. The annual gross electricity generation from coal has increased from 612497 GWh in year 2011-12 to 995840 GWh in year 2019-20 while generation from solar energy has increased to 50103 GWh in 2019-20 [35-37]. In coming years, the power generate through solar energy will increase with high rate.

Table 1 Annual gross electricity generation from conventional energy and RECL economic support in GWh [31-37]

	Conventional Energy (GWh)			**REC (GWh)**					
Year	**Coal**	**Oil**	**Gas**	**Hydro**	**Solar**	**Wind**	**Biomass**	**Other**	**Total**
2019-20	995,840	108	48,497	9,366	50,103	64,639	13,843	366	138,318
2018-19	1,021,997	129	49,886	8,703	39,268	62,036	16,325	425	126,757
2017-18	986,591	386	50,208	5,056	25,871	52,666	15,252	358	101,839
2016-17	944,850	262	49,100	7,673	12,086	46,011	14,159	213	81,949
2015-16	896,260	406	47,122	8,355	7,450	28,604	16,681	269	65,781
2014-15	835,838	1,407	41,075	8,060	4,600	28,214	14,944	414	61,780
2013-14	746,087	1,868	44,522	0	3,350	0	0	0	59,615
2012-13	691,341	2,449	66,664	0	0	0	0	0	57,449
2011-12	612,497	2,649	93,281	0	0	0	0	0	51,226

4. Energy Requirement and Supply

4.1. Energy Requirement

India has a large population therefore energy demand is also very high. The Indian energy consumption market contributes about 5.6 % of the total world's energy consumption. Coal contributes about 56.90% of total energy generation in India, while oil contribution is 29% [38]. The energy consumption in the industry sector is the highest, which is almost half of the total generated energy, followed by the transportation division ranks second which is about 22% [39]. Considering Indian-household division, traditional biomass contributes about 33.33% of net energy required. 30 % of the total is spent in lighting while only 1% is used up for cooking. The

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

remaining energy, which is about 40%, is spent in home-related devices [40]. The agricultural sector consumed 18% of the total produced electricity.

The country's electric energy division today requires more than 40% of basic power asset, with an enormous input from hydroelectricity (25%), coal (70%), trailed by natural gas, renewable assets, and nuclear energy responsible for the left 5% [41-42].

Table 2 depicts the energy demand and peak demand in 2009-10 was -83950 MW deficit and 15158 surplus respectively. This energy demand increased to -940694 MW deficit while peak demand reduced -1147 MW deficit in 2020-21 [43-45].

Power generation climbed by 5.71 percent from 2007-08 to 2017-18, whereas electricity consumption increased by 7.39 percent. Furthermore, the state of electricity distribution in India from 2008-09 to 2017-18 demonstrates that, while transmission and distribution (T and D) losses in India have declined from 25.7 percent in 2008-09 to 21.15 percent in 2017-18, the decrease is not significant [46]. In the last ten years, the energy requirement has increased at a rate that has outpaced the expansion of the Gross Domestic Product (GDP) [47]. Electricity consumption per person increased to 1,149 KWh in 2017-18, compared to around 90 KWh in 1971-72 [48].

As a result, the demand for energy, GDP, and people is expected to skyrocket in the next twenty years, as shown in Fig. 4. According to Yang et al. [49], as the population grows, demand for net and total energy increases linearly.

Table 2 Annual Energy requirement in India from 2009-10 to 2020-21 [43-45]

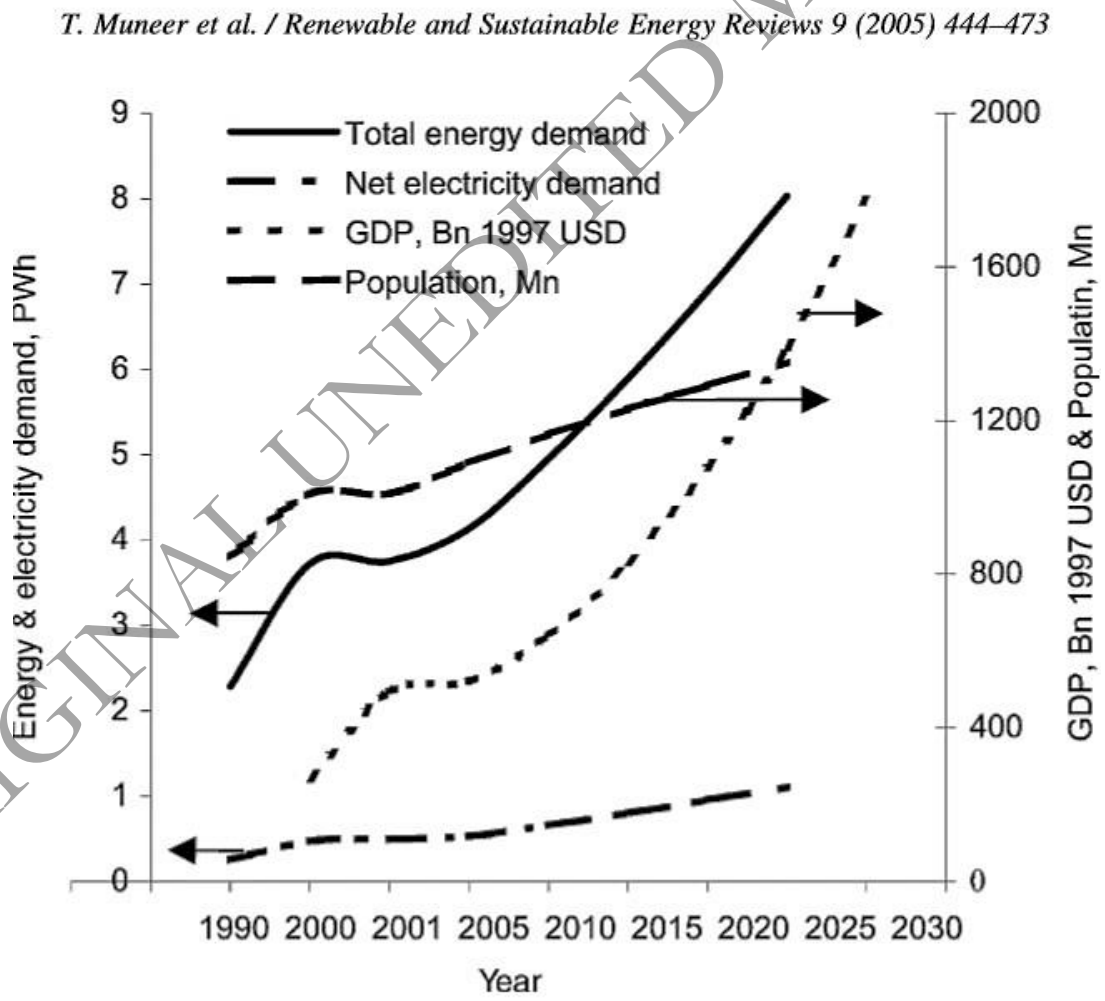
Energy Requirement				**Peak Demand**				
Year	**Demand**	**Availability**	**Surplus (+)/ Deficits (-)**	**Peak Demand**	**Peak Met**		**Surplus (+) / Deficits (-)**	
	Mega unit (MU)	**(MU)**	**(MU)**	**(%)**	**(MW)**	**(MW)**	**(MW)**	**(%)**
2009-10	8,30,595	7,46,645	-83,950	-10	1,19,170	1,04,010	15,158	-12.8
2010-11	8,61,590	7,88,356	-73,238	-8.3	1,22,288	1,10,258	12,030	-9.7
2011-12	9,37,200	8,57,888	-79,316	-8.8	1,30,007	1,16,193	13,814	-10.5
2012-13	9,95,560	9,08,652	-86,904	-8.4	1,35,455	1,23,295	12,160	-9.1
2013-14	10,02,259	9,59,829	-42,430	-4.0	1,35,920	1,29,816	-6,100	-4.6

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

2014-15	10,68,919	10,30,785	-38,140	-3.4	1,48,164	1,41,161	-7,005	-4.8
2015-16	11,14,408	10,90,850	-23,560	-2.3	1,53,363	1,48,462	-4,905	-3.3
2016-17	11,42,929	11,35,334	-7,596	-0.9	1,59,540	1,56,935	-2,609	-1.7
2017-18	12,13,326	12,04,697	-8,630	-0.9	1,64,067	1,60,753	-3,317	-2.2
2018-19	12,74,595	12,67,526	-7,069	-0.7	1,77,023	1,75,530	-1,495	-0.9
2019-20	12,91,010	12,84,444	-6,559	-0.4	1,83,805	1,82,535	-1,270	-0.8
2020-21	9,40,694	9,37,518	-3,176	-0.3	1,84,033	1,82,888	-1147	-0.7

Fig. 4 Economical energy and population trend in India [49]



ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

4.2. Energy Supply

4.2.1. Conventional Energy

The conventional energy sources are non-renewable energy sources. In the current environment, coal and petroleum are the most frequently used nonrenewable energy resources for energy generation. Natural resources are depleted because of their widespread use. In India alone, known petroleum deposits are expected to last only a few decades, whereas coal reserves are expected to last barely another hundred years. The rising energy cost has compelled world governments to recognize the need for the research into alternate energy sources.

(a) Coal and Lignite-

The major coal producing states are Jharkhand, Madhya Pradesh, Orissa, Maharashtra, Andhra Pradesh, etc. Table 3 shows the overall state wise coal reserves in India. Among them, Jharkhand ranks first in coal production (26.2%) while Orissa ranks second with a contribution of 25.1 % . Sikkim, Arunachal Pradesh accounts the lowest distribution 0.3% and 0.2% respectively (table 3).[50]. The net amount of coal reserves estimated to be 319.04 billion tons while that of lignite were 43.24 billion in 2019-20 [51].

Table 3. State wise overall coal reserves in India [50-51]

States	Total coal reserves as on 31.03.2020 (BT)	Distribution (%)
Jharkhand	83.13	26.2
Odisha	79	25.1
Chhattisgarh	57.24	17.9
West Bengal	31.68	9.9
Madhya Pradesh	28	8.6
Telangana	21.8	6.5
Maharashtra	12.4	3.8
Andhra Pradesh	1.6	0.49
Bihar	1.4	0.42
Uttar Pradesh	1.0	0.33
Meghalaya	0.48	0.18
Assam	0.43	0.15
Nagaland	0.40	0.13
Sikkim	0.09	0.03
Arunachal Pradesh	0.09	0.02
Total	**319.04**	**100**

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

(b) Petroleum and Natural gas

The assets of crude oil in our country were estimated as 762.7 metric tons (MT) in March 2020 [52]. The maximum reserves of crude oil are in offshore (Western) (42.91%) trailed by Assam (22.69%). Table 4 shows the region wise distribution of crude oil in Andhra Pradesh (1.5%). Eastern offshore (7.35%), Gujarat (17.7%) and Eastern offshore (42.9%).

On the other hand, the assets of Natural Gas are more in Offshore (Eastern) (37.34%) trailed by Western offshore (30.17%) [53-55]. The region wise distribution of natural gas in Andhra Pradesh, offshore (Eastern), Gujrat, and Western offshore is 3.38 %, 37.24 %, 5.11 %, and 30.17 % respectively (Table 4) [57].

Table 4 Crude Oil and Natural Gas estimates and distribution in India as of March 2020 [55-58].

States/ Region	**Crude Oil (MT)**		**Natural Gas (Billion Cubic Meters)**	
	Estimated reserves	Distribution (%)	Estimated reserves	Distribution (%)
Andhra Pradesh	11.45	1.50	48.20	3.38
Eastern Offshore	56.09	7.35	531.54	37.24
Gujarat	135.01	17.70	72.96	5.11
Western Offshore	327.28	42.91	430.61	30.17

The contribution of conventional asset in energy generation estimated to be 215000 GWh majorly dependent on coal accounting to 186577 GWh from Table 5 [63].

4.2.2. Renewable assets contribution

Renewable assets are presently contributing about 38 % of the total energy supply in India for the year 2020 [59]. India aimed to generate 275 GW of electricity from renewable energy by 2027 [60]. Thus, the growth rate of renewable energy-based power generation is much higher than any other power source. It is also predicted that the contribution of renewable energy-based power generation will go up to 70% by the year 2051[61]. If the current situation continues, the world will have to finance a total of 1000 lakh crore Indian rupees for the next 30 years [60]. India has set a goal of producing 175 GW of renewable electricity by 2022 [62].

Among them, solar power is one of the most dependable and predictable assets. Table 6 presents the renewable contribution share in India. The total installed capacity of renewable asset is

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

89229 MW alone solar contribute about 36051 MW [64]. The generation of energy has limitations in relation to the type and availability of land. While, on the other hand, the utilization of solar energy does not concern these problems [65]. The use of solar photovoltaic energy is encouraged more as they are considerably cheaper than the capital required for the way we generate electricity today from fossil fuels. It has various applications, for example, the use of nanoscale-antennas to convert 80-90% of sunlight into energy [66].

Table 5 Energy consumed through fuel as of March 2020 [63].

Fuel	Electricity generated (GWh)	Total share
Hydro	280.12	0.09 %
Nuclear	0.035	0.01 %
Natural Gas	19,472.59	9.06 %
Oil	4818.87	2.24 %
Coal	186,577.73	86.78 %
Total	215,000	100.0

Table 6 Renewable asset contribution as of March 2020 [64].

Renewable power capacity			**Grid connected installed capacity from sources**		
Source	Total Installed Capacity (MW)	2022 target (MW)	Source	Installed Capacity (MW)	Share (%)
Wind power	38,124	60,000	Coal	205,135	55.43
Solar power	36,051	100,000	Large hydro	45,699	12.35
Biomass power			Other renewables	87,028	23.51
(Biomass & Gasification and Bagasse Cogeneration)	10,146	10,000	Gas	24,955	6.74
Waste-to-Power	169	10,000	Diesel	510	0.14
Small hydropower	4,740	5,000	Nuclear	6,780	1.83
Total	**89,229**	**175,000**	**Total**	**370,106**	**100**

5. Renewable Energy Utilization in India

These assets can be generated through various sources- solar, wind, hydro, biomass, etc. Their potential as per source in Table 7 and different values inferred from the given Table 7. The overall renewable energy installation as of 2020 is estimated to be 85908 MW while the target

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

for 2022 is 27000 MW [67]. Gujarat (about 25%) ranks first in the estimated potential of renewable assets as per states, trailed by Karnataka and Tamil Nadu [67].

Table 7 Renewable Energy Installations as of March 2020 [67].

Sector	Cumulative Achievements (as of 2020) (MW)	2022 Target (MW)
Wind	37505	175,000
Solar Power - Ground Mounted	31380	10,000
Solar Power - Ground Mounted	2334	60,000
Small Hydro Power	4671.6	5,000
Bio Power (Biomass & Gasification and Bagasse Cogeneration)	9860.98	10,000
Waste to Power	139.80	10,000
Total	**85,908.37**	**270,000**

5.1 Wind Energy:

Wind energy is a renewable asset much dependable during recent years. Its use in India has increased. Its contribution alone is 68% of the total renewable asset in India. As per the Centre for Wind Energy Technology (C-WET), at height of 80 meters with about 2% land availability, wind potential would be over 100GW [68]. The 12th Five Year Plan targets installing about 15,000 MW in the span (2012-17), making it twice the current power generation [69]. The state-wise wind energy capacity and top 10 Indian states with maximum wind energy production are estimated in Table 8 – 9. Table 8 depicts Tamil Nadu has the highest wind energy capacity accounting of 11900 MW and 24.2 % share [70]. Table 9 shows some of the important windmills in India is Muppandal Wind Farm, Jaisalmer Wind Park, Brahmanvel windfarm, Kayathar and Dhalgaon windfarm [71].

Table 8 State wise yearly wind energy capacity [70].

S. No	State	2014 (MW)	2015 (MW)	2016 (MW)	2017 (MW)	2018 (MW)	2019 (MW)	2020 (MW)	% Share
1	Tamil Nadu	7276	7515	7694	7970	8435	9231.77	11,900	24.2
2	Maharashtra	4098	4638	4666	4778	4990	4794.13	7600	14.5

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

3	Gujarat	3414	3877	4441	5537	5789	7203.77	8800	16.8
4	Rajasthan	2820	3866	4217	4282	4698	4299.73	8600	13.0
5	Karnataka	2409	2872	3154	3793	4198	4753.40	6200	11.5
6	Andhra Pradesh	753	1155	2092	3835	4098	4077.37	8100	11.6
7	Madhya Pradesh	439	1126	2288	2498	2786	2519.89	6200	7.6
8	Kerala	55	35	43	51	65	62.50	–	0.15
9	Telangana	-	–	99	101	123	128.10	2000	0.31
10	Others	-	4	4.30	4.30	4.30	4.30	600	0.013
Total		**21,264**	**25,088**	**28,700**	**32,848**	**35,186.3**	**37,090.03**	**60000**	**100**

Table 9 Top 10 wind energy production in India [71-72].

Rank	Power Plant	Producer	Location	State	MW
1	Muppandal Wind Farm	Muppandal Wind	Kanyakumari	Tamil Nadu	1500
2	Jaisalmer Wind Park	Suzlon Energy	Jaisalmer	Rajasthan	1063
3	Brahmanvel windfarm	Parakh Agro Industries	Dhule	Maharashtra	530
4	Kayathar	Siemens Gamesa, ReNew Power	Tutcorin	Tamilnadu	303
5	Dhalgaon windfarm	Gadre Marine Exports	Sangli	Maharashtra	280
6	Vankusawade Wind Park	Suzlon Energy Ltd.	Satara District.	Maharashtra	260
7	Vaspet	ReNew Power	Vaspet	Maharashtra	143
8	Tuljapur	Siemens Gamesa, ReNew Power	Osmanabad	Maharashtra	127
9	Beluguppa Wind Park	Orange Renewable	Beluguppa	Andhra Pradesh	99.8
10	Mamatkheda Wind Park	Orange Renewable	Mamatkheda	Madhya Pradesh	99.5

5.2 Small Hydropower (SHP):

Hydro projects were generating electricity less than 25 MW known as Small Hydropower [73]. They are the oldest means of renewable assets for generating power. Table 10 shows state wise net installed hydropower plant capacity in India. The maximum yield installed capacity today is less than 2500 MW in Uttarakhand. The 12th Five Year Plan targets to bring an increase of 2,100 MW in SHP [74].

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Table 10 Hydropower plant project in India [74].

Name	River	Location	Capacity (MW)
Tehri Dam	Bhagirathi	Uttarakhand	2400
Koyna	Koyna	Maharashtra	1960
Srisailam Dam	Krishna	Andhra Pradesh	1670
Nathpa Jhakri	Satluj	Himachal Pradesh	1500
Sardar Sarovar	Narmada	Gujarat	1450
Bhakra Dam	Satluj	Punjab	1325
Sharavathi	Sharavati	Karnataka	1035
Indira Sagar	Narmada	Madhya Pradesh	1000
Pandoh Dam	Beas	Himachal Pradesh	990
Nagarjunasagar	Krishna	Andhra Pradesh	965
Kalinadi	Kalinadi	Karnataka	955
Idukki	Periyar	Kerala	780
Salal	Chenab	Jammu & Kashmir	690
Loharinag Pala	Bhagirathi	Uttarakhand	600
Omkareshwar	Narmada	Madhya Pradesh	520
Teesta VI	Teesta	Sikkim	510
Loktak	Manipur	Manipur	105

5.3 Biogas and biomass energy

Biomass is a renewable extracted from the carbonaceous left-over. It is obtained from sources like timber, animal dung, agricultural residues such as bagasse, and wastes generated from agro-based industries. Biomass receives carbon from the atmosphere and returns in burned form. It is biodegradable in nature. Solid wastes, animal wastes, etc., are considered as biomass. Table 11 shows the estimated biomass installed capacity in Indian states. The maximum energy potential is in Punjab, having 2414 MW, while Uttar Pradesh have a maximum installed capacity of 592.6 MW [75].

Table 11 Installed biomass capacity in Indian States [75]

State	Energy Potential (MW)	Installed Capacity (till 2020) (MW)	Tariff rate per unit (INR)
Punjab	2414	74.61	5.25
Uttar Pradesh	1594.3	592.6	4.7
Haryana	1120.9	35.9	5.25
Rajasthan	1093.6	73.2	4.7

ORIGINAL UNEDITED MANUSCRIPT

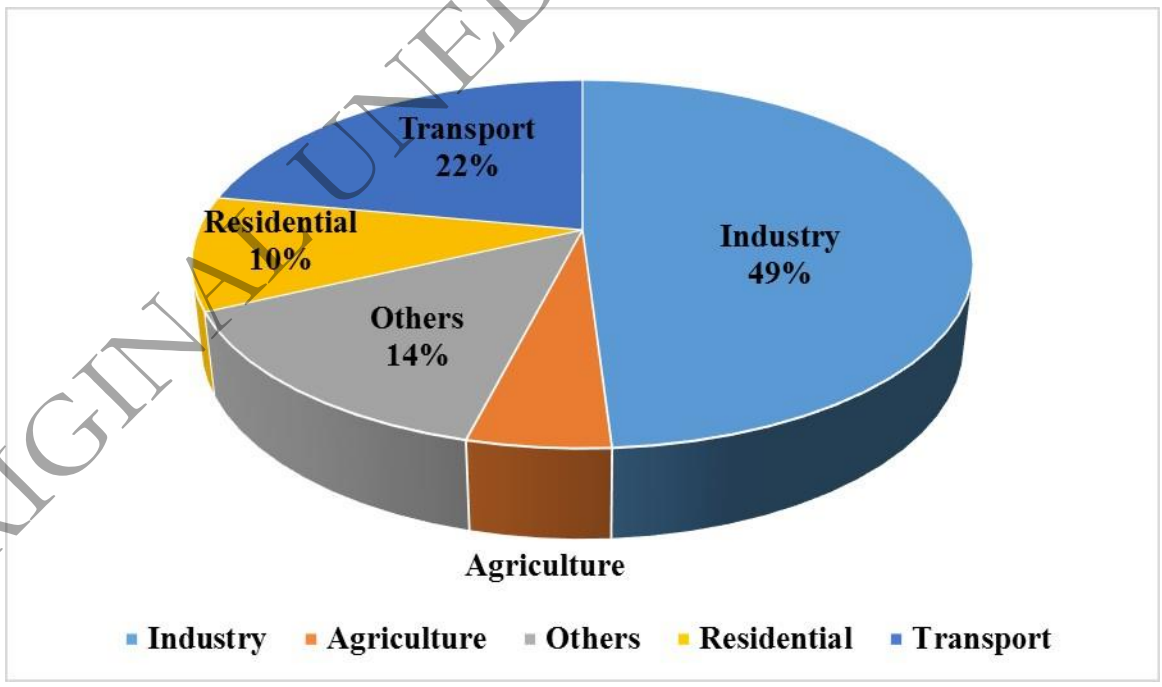
Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Maharashtra	1014.3	402	5
Madhya Pradesh	842	1.0	4-5.14
Karnataka	632	365.2	4-5
Andhra Pradesh	624	363.3	4.5 t
Gujarat	458	0.6	4.8
Chhattisgarh	248.6	232	4
Kerala	196	–	3

5.4 Solar Energy:

Solar Energy is used not only for heating or thermal processes but also for the generation of electricity. The net amount of solar energy we receive is around the order of 10^{16} MW, which is a very high value [76]. This value is 1000 times more than the world's net energy requirement. As per MNRE, India has generated more than 2 GW power from solar energy [77]. Thus, Renewable assets have a lot of scope and advantages. Fig. 5 shows the Indian sectoral composition of commercial energy consumption. The total energy consumption shares in Industry, transportation, residential, agriculture and other sectors are 49 %, 22 %, 10 %, 4 % and 14 % (Fig.5) [78].

Fig. 5 Share of energy consumption in India [78]



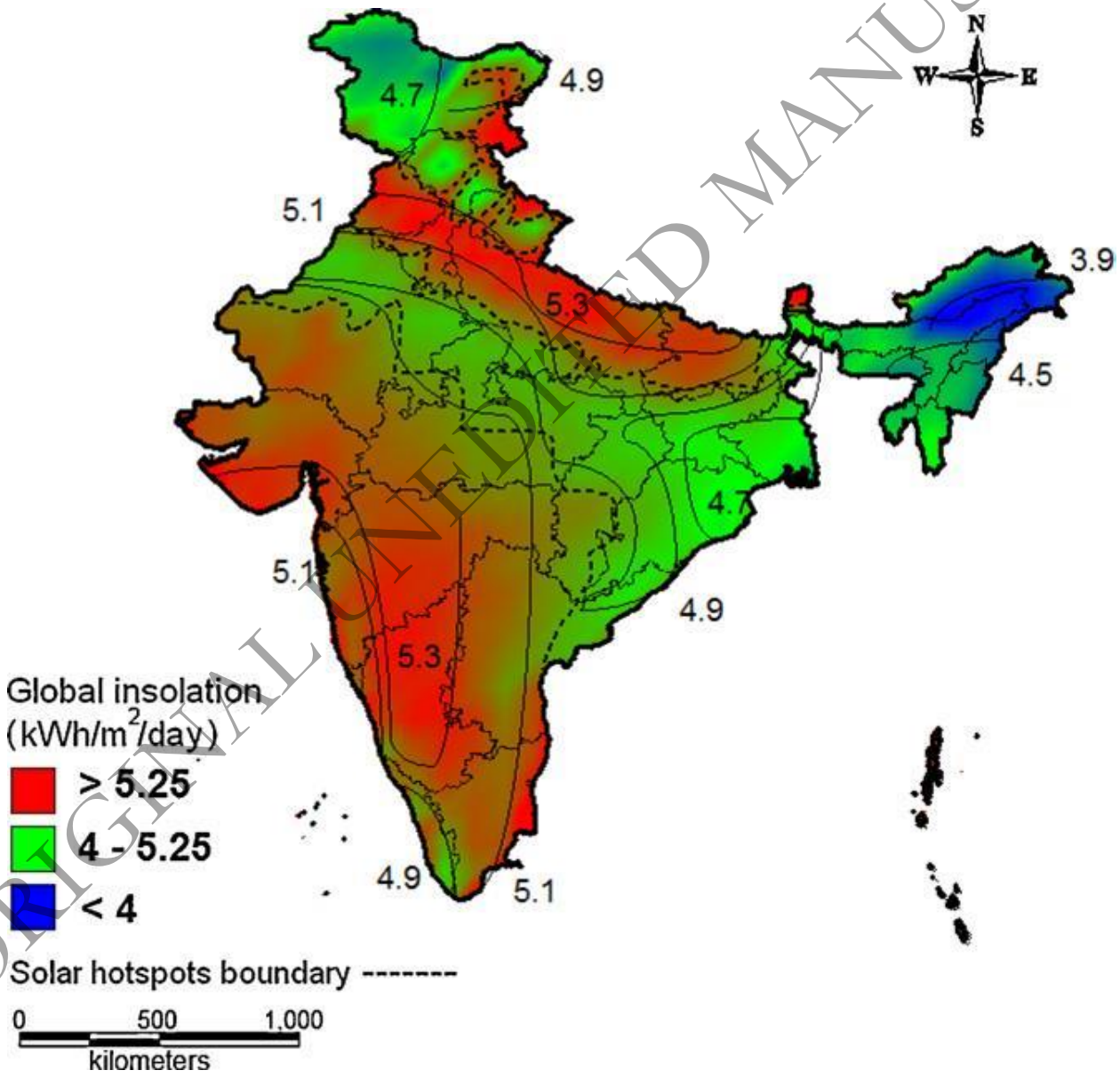
ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

6. Solar Energy Resources in India

As discussed earlier, solar energy has a lot of scopes to be used in the Indian power generation market. A new keyword, "Solar Hotspot" has been introduced to know the potential of regions for electricity generation using solar energy. As per data collected from different sources, it has been found that almost 58 % of Indian land can be considered as solar hotspots, receiving energy of about 5 KWh/m^2 per day [80]. Fig. 6 represents the annual global solar radiation for India.

Fig. 6 India annual global solar irradiation distribution in KW h/m^2 [79].



ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Figure 6 indicates the maximum solar irradiation received is in Jodhpur (Rajasthan), which is estimated to be 2173 KW h/m^2 and minimum in Calcutta (West Bengal) 1805 KW h/m^2 [79].

Figure 7 shows the majority of renewable energy potential is from solar energy, accounting for 40 % [81]. The solar energy potential in different Indian states is estimated in Fig. 8 and Rajasthan, share maximum solar potential accounting to 21% as compared to other Indian states [82]. As per the economic perspective, not only solar potential considered and the economic organizational is important. Therefore, the region with all three common factors, including techno economical factor, solar potential region, and organizational infrastructure, is considered as solar hotspot. This diagrammatic representation of the solar hotspots represented in Fig. 9. The amount of solar energy received depends on the nature of suspended particles in the atmosphere, the tilt of the earth on its axis, etc. The "Solar Belt" is a region on earth from (40˚S to 40°N) India comes under this region [83]. Thus, solar energy received is being utilized using Solar PV, solar lanterns, solar pumps, solar heaters, solar thermal water heaters, solar cookers, household lighting/cooking systems, etc., which aggregates to 66 MW power [84]. Among which around 12 MW power grid-connected while 2 MW off-grid power connected [84]. National Solar Mission (NSM) aims to achieve 22,000 MW power through grid connection while 2,000 MW power as off-grid connection by the year 2022, done possibly by discovering hotspots which have a great potential to generate solar power considering global insolation (per day) the characteristics are as follows:

I. Southern Peninsula-4.51 kWh/m^2

II. Western and eastern Himalayas- 3-4 kWh/m^2

III. During summer, more than 90% of India receives more than 5.5 kWh/$m^{2.}$

IV. Western dry regions and central plateaus -4.69 kWh/m^2

V. Western Himalayas in Northern regions of India-2.5kWh/m^2

Solar Energy Applications for Power Generation are as follows:

I. Solar energy has wide applications, and one of them is the use of solar energy for power generation.

II. There are several ways in which solar energy can be utilized to generate electricity. Some of them are mentioned below:

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Fig. 7 Source wise estimated Renewable Energy potential in India (March 2020) [81].

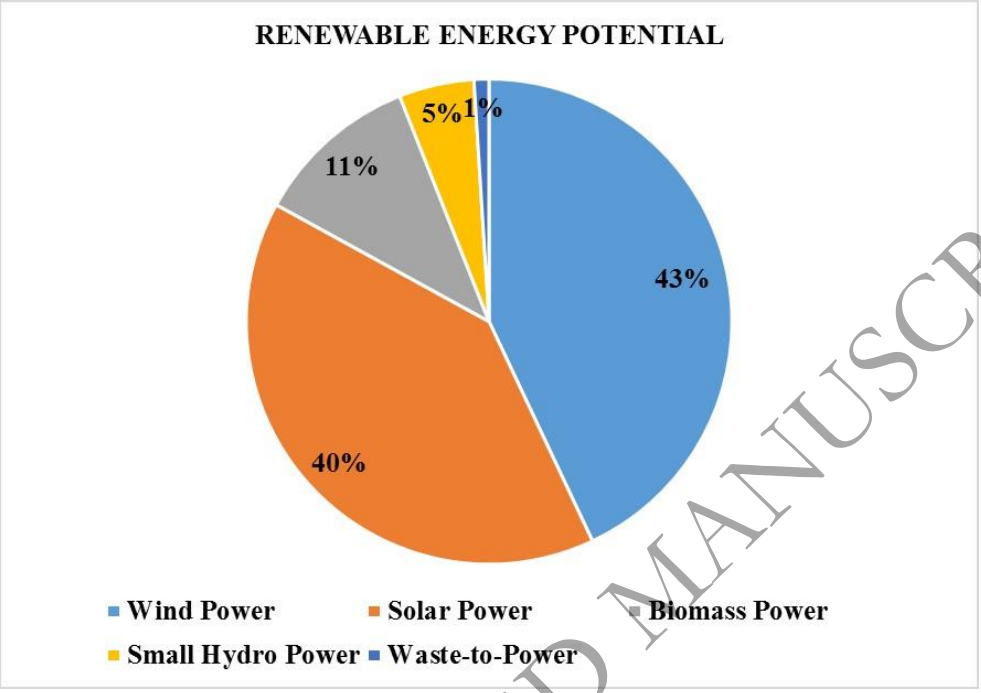
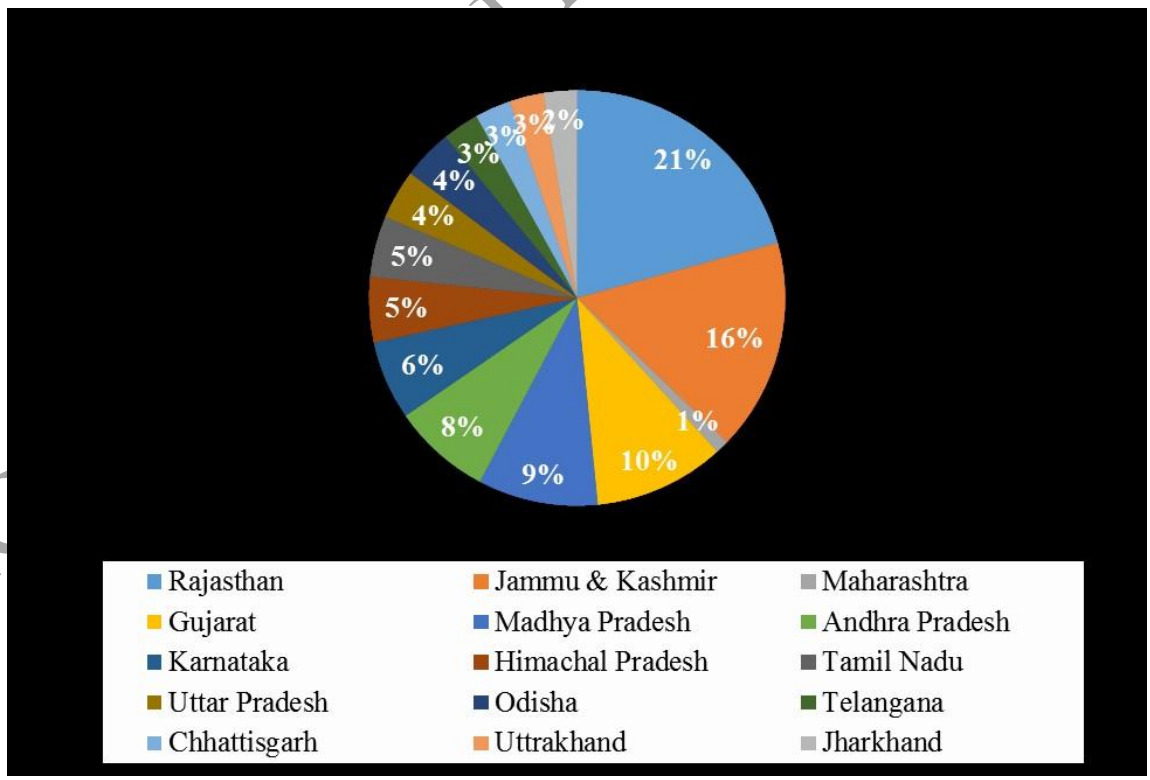


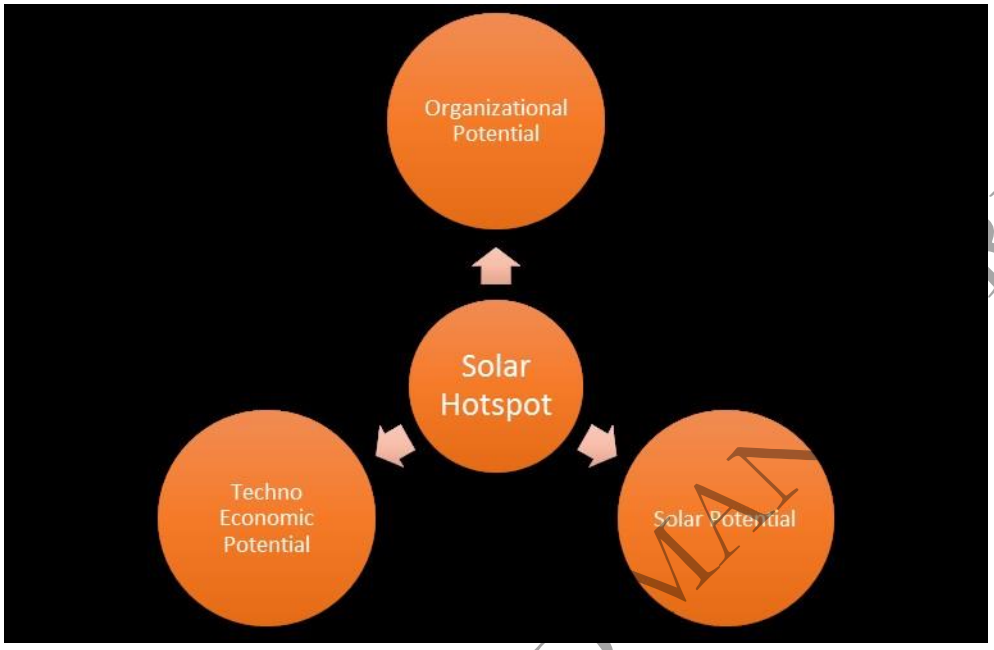
Fig. 8 Estimated Renewable Energy potential in States of India (March 2020) [82].



ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Fig. 9 Solar hotspots representation [83].



6.1 Solar PV Energy

Solar PV (Photovoltaic) energy or SPV (Solar Rooftop Photovoltaic) technic is based on the principle of conversion of sun radiation directly through devices made of semi-conductor to generate electric power [85]. These devices are known as solar cells. Such cells are connected to each other to form a module that is further interconnected to give a solar battery that may be used as per energy requirement. These modules have a scope of high usage soon, as they are more reliable with better efficiency. PV cells are used in urban residence electricity requirement, and spacecraft. Table 12 depicts India's major PV power plants and mentioned few like Bhadla Solar Park, Galiveedu solar park, NP Kunta, etc. [86]. It can also be used as a substitute for coal if the region is a good hotspot or has good solar power generation potential. Therefore, to harness solar energy majority of dependency lie on solar PV cell. Thus, solar PV capacity has increased globally with the moving years depicted in Fig.10 [87]. However, India is yet to match the world annual growth PV module shipment, and the red line represents the global PV capacity while the black line represents India's PV capacity [88]. The current market of SPV technique divided into three divisions:

I. Professional sector- By usage of repeater telecommunications system, which are remotely controlled.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

II. Grid-connected sector- Supplies sufficient power back to the utility system received from the provision of homes having a low electric power supply.

III. Off-grid sector- Manufactures power independently.

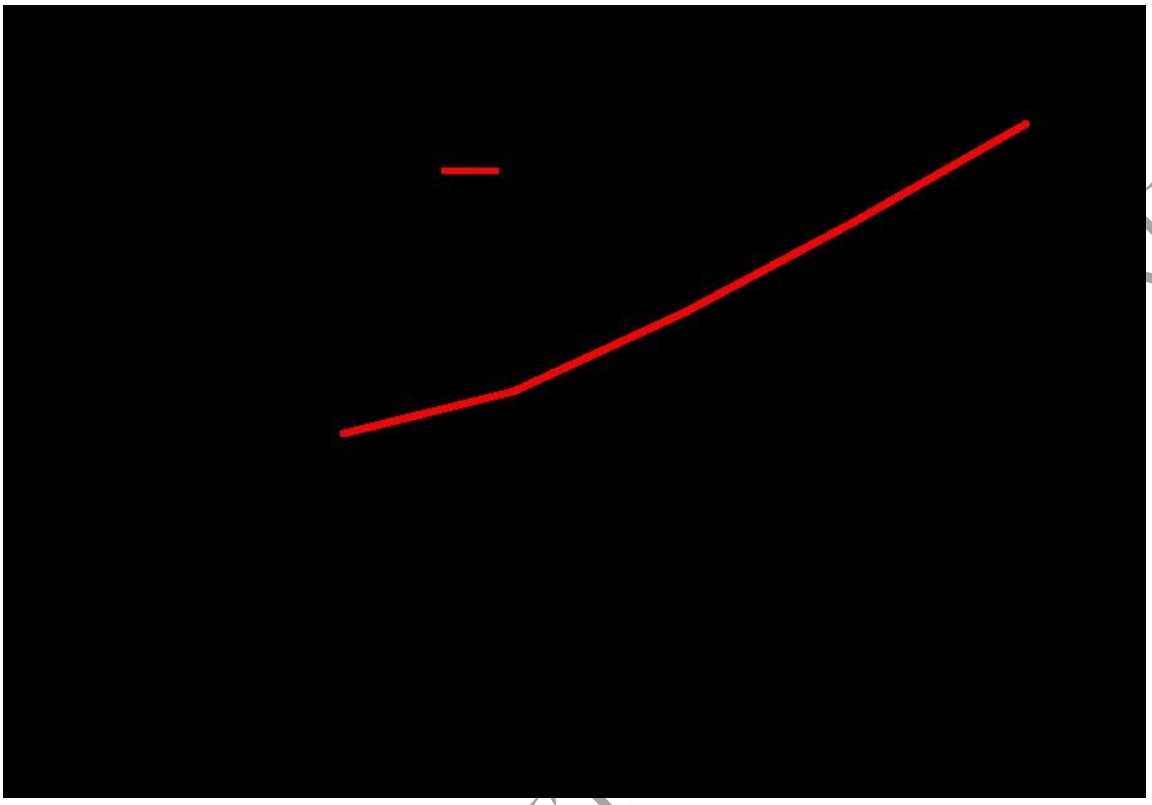
Table 12 Major photovoltaic (PV) power plants [86].

Plant	State	DC peak power (MW)	Commissioned Year
Bhadla Solar Park	Rajasthan	2245	2020
Galiveedu solar park	Andhra Pradesh	400	2020
NP Kunta	Andhra Pradesh	900	2020
Kadapa Ultra Mega Solar Park	Andhra Pradesh	250	2020
Pavagada Solar Park	Karnataka	2,050	2019
Ananthapuramu	Andhra Pradesh	400	2019
Karnataka I	Karnataka	40	2018
ReNew Power, Nizamabad	Telangana	143	2017
Kamuthi Solar Power Project	Tamil Nadu	648	2017
Mandsaur Solar Farm	Madhya Pradesh	250	2017
Maharashtra I	Maharashtra	67	2017
Jalaun Solar Power Project	Uttar Pradesh	50	2016
NTPC solar plants	UP	110	2015
Welspun Solar MP project	Madhya Pradesh	151	2014
Green Energy Development Corporation (GEDCOL)	Odisha	50	2014
Tata Power Solar Systems (TPS), Rajgarh	Madhya Pradesh	50	2014
GEDCOL	Odisha	48	2014
Sakri solar plant	Maharashtra	125	2013
Welspun Energy, Phalodhi	Rajasthan	50	2013
Charanka Solar Park	Gujarat	690	2012
Gujarat Solar Park	Gujarat	221	2012
Bitta Solar Power Plant	Gujarat	40	2012
Dhirubhai Ambani Solar Park, Pokhran	Rajasthan	40	2012

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Fig. 10 Annual growth in PV module shipment [87].



6.2.1 Using concentrating solar power by concentrating collectors (CSP)

The type of collectors used nowadays to generate steam to heat water for domestic purposes or to generate steam to run turbine coupled with a generator to generate electric power. Some types are flat plate collectors, parabolic collectors (which can be both uni-axial and bi-axial), etc. [89]. Parabolic collectors are much encouraged nowadays to concentrate solar power. Presently, euro trough parabolic collectors have gained much attention because of their high performance and efficiency. Their operating temperatures can be as high as 400-500 °C. Fig. 11 represents euro trough collector element consisting of endplates, steel frames screwed to a torque box, absorber tube support, cantilever arms and 28 mirror facets while Table 13 represent a parameter related to euro trough 100 m and 150.8 m [90-91]. The computer-aided modelling of Euro trough collector with torque box design and concentrator deformation analysis was carried out in Fig. 12 and 13 [92]. The different concentrator design under bending and torsion load modelling done in Fig. 13. Fig. 13 depicts the regular concentrator, torque tube and torque box design concentrator modelling done under different support structure [93].

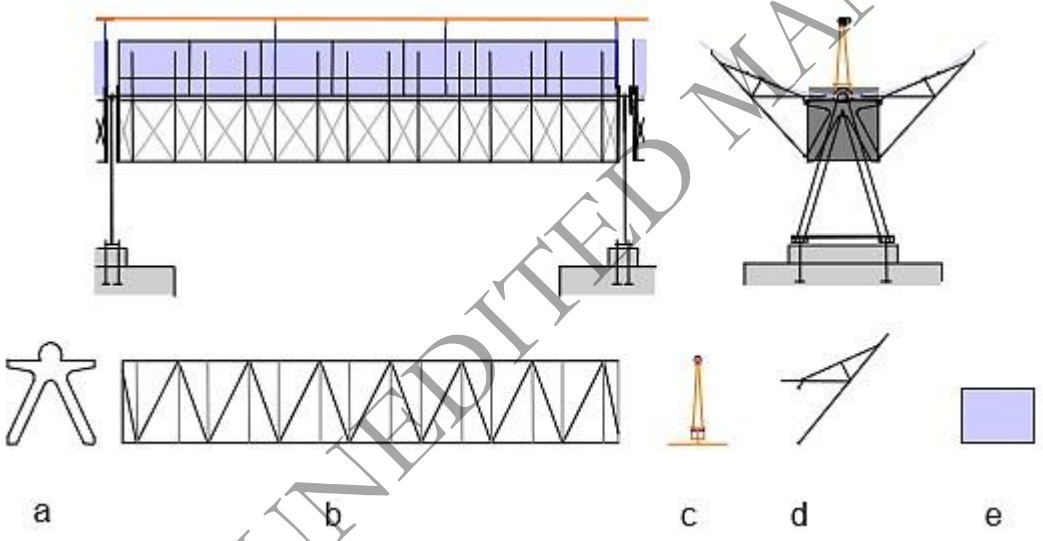
ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Table 13 Parameters related to Euro trough 100 m and 150.8 [90]

Euro trough Model	ET 100	ET 150
Aperture Area	544m^2	818m^2
Focal length	1.69m	1.69m
Absorbing Radius	3.5 cm	3.5 cm
No. of glass facets	225	340
Mirror reflectivity	93.8%	93.9%
Aperture width	5.77 m	5.77 m

Fig. 11 Euro trough collector element consisting out of **a**) 2 end plates; **b**) 4 simple steel frames screwed to a torque box; **c**) 3 absorber tube supports; **d**) 28 cantilever arms **e**) 28 mirror facets [91].



ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Fig. 12. Computer model of the Euro trough collector with torque-box design [92].

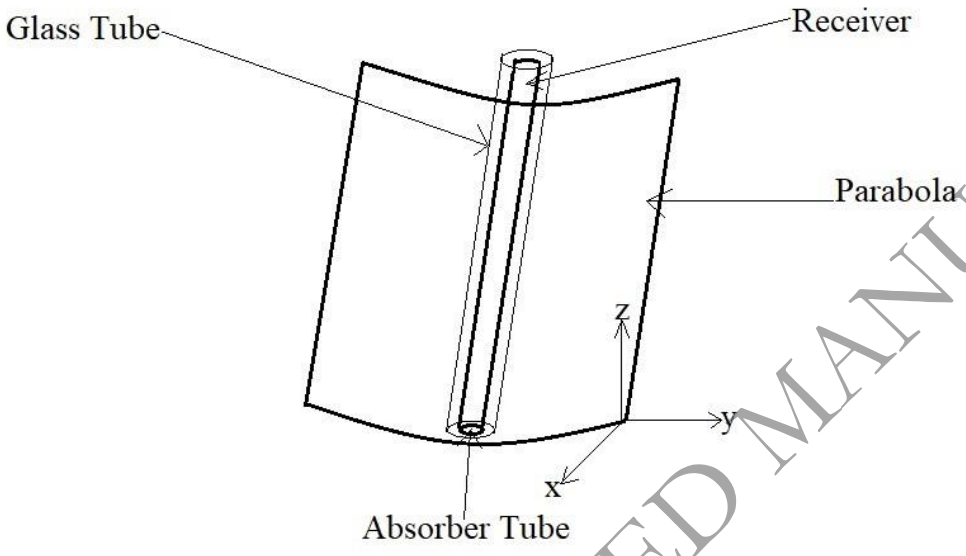
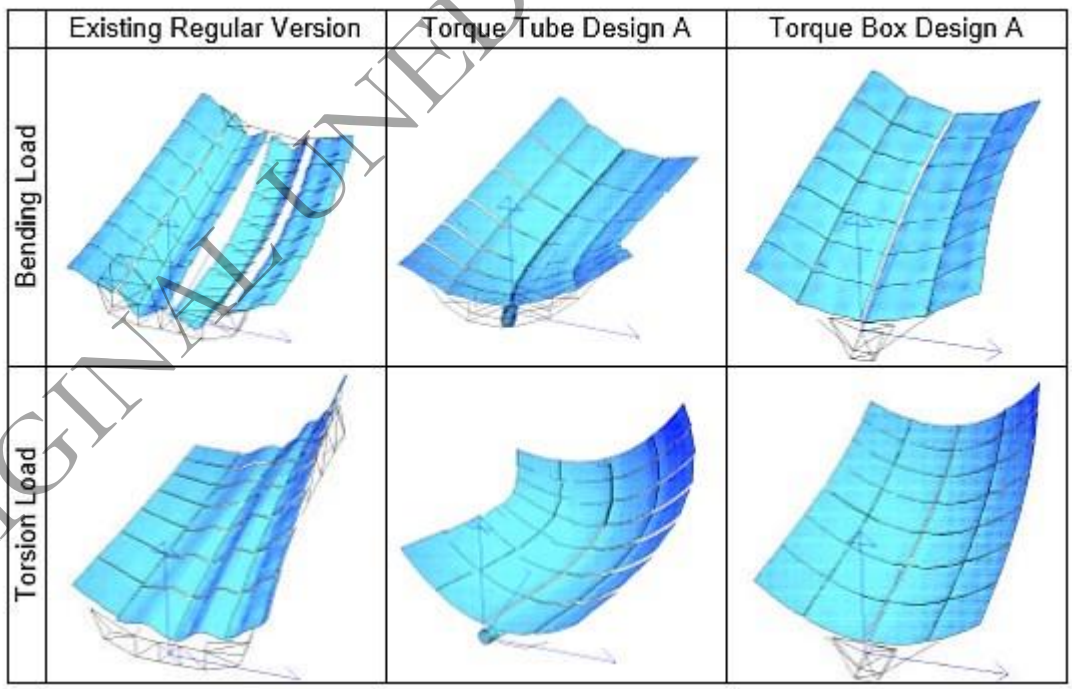


Fig. 13. Concentrator deformation analysis for different support structure [93].



ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

6.2.2 Working Principle

The parabolic shape of collectors enables the radiation coming from the sun to concentrate on the focal line where the absorber tubes are kept. A heat transfer fluid (HTF) flows through these absorber tubes, which is heated nearly up to 400°C [94].

Other applications include:

I. Rural Electrification
II. Solar lamps & lighting
III. Agro-support
IV. Solar Water Heaters
V. Power Grids

Table 14 shows the state-wise top ten cumulative solar-capacity.

Table 14. State wise top 10 Cumulative Solar Capacity [96-97].

State/UT	Cumulative Solar Capacity until 31-03-19 (MW)	Added solar capacity in 2019-20 until 31-12-19 (MW)	Cumulative Solar Capacity until 31-12-19 (MW)
Karnataka	6095.55	1179.37	7274.92
Rajasthan	3226.79	1617.42	4844.21
Tamil Nadu	2574.98	1212.99	3787.97
Telangana	3593	28.70	3621.00
Andhra Pradesh	3085.68	473.34	3559.02
Gujarat	2440.13	323.42	2763.55
Madhya Pradesh	1839.9	398.00	2238.50
Maharashtra	1634.12	30.10	1662.98
Uttar Pradesh	960.10	85.00	1045.10
Punjab	905.62	41.48	947.10

6.3 Manufacturing/assembling facilities

There are many facilities or companies which nowadays manufacture solar appliances, device, and equipment.

I. In Andhra Pradesh, APTrans Company has made agreements with IPPs to install about 1000 MW worth solar power plant, which also includes the installation of 250 MW in Anantpur district

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

II. Gujarat share the maximum production of solar power in the whole country, about 66.66%. It is planned to gain about 500 MW of solar power from it.

III. The French AREVA group has been engaged with the Rajasthan to install CSP Plant worth 250 MW, which would be the largest in Asia [94].

IV. An Ultra Mega Green Solar Power Project worth 4,000 MW is being installed near sambhar-lake.

V. Sakari Solar power plant worth 125 MW is the biggest in Maharashtra. Shri BABA SANTHAN TRUST is most extensive solar steam system on the entire earth and is situated in Maharaahtra [95].

VI. Ujaas Energy Limited and Welspun Solar MP projects in Madhya Pradesh belong to the world's top solar power generation plants.

Table 15 exhibits the state-wise capacity of Solar rooftop systems.

Table 15. State-wise Solar Rooftop Systems Capacity (as of March 2020) [98]

State /Uts	Subsidized (MW)	Non-Subsidized (MW)	Total (MW)
Gujarat	220.83	80.88	301.71
Haryana	57.76	60.70	118.47
Karnataka	23.70	108.12	131.83
Maharashtra	95.69	120.42	216.11
NCT Of Delhi	95.33	14.47	109.80
Rajasthan	59.54	59.96	119.50
Tamil Nadu	60.53	95.25	155.78
Uttar Pradesh	89.69	51.18	140.87

7. Government Policy and Infrastructure for Promoting Solar Energy

Indian federal consists of 28 states and 08 union territories (UT). Each state and UT have its own energy policy, which is quite supportive of the central government policy. The Indian government is very keen to increase the total capacity of solar power; therefore, regular upgradation of the approach related to solar energy happens. Table 16 presents in brief form the various important policies of the central government from 1987 to till now [99]. The government provides easy loans, subsidies for both consumers and entrepreneurs, land for the plant and relaxation in income and sale tax.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Table 16 Central government solar policy [99].

Year	Policy	Aim	Status
1987	Indian Renewable Energy Agency (IREDA)	Establishment as a public financing organization. Provide financial support new and renewable sources and conserving energy through energy efficiency.	Improvement in efficiency of services provided to customers
1993	The Electricity Laws Act	Income tax exemption Tax holiday Excise and custom duty benefit	Amendment of electricity act 2003
2004	Renewable Purchase Obligation (RPO) policy launched in Maharashtra	The power generation through the renewable energy sources every year declared by the state and central government	Announced National Tariff Policy in 2006
2009	Generation Based Incentives (GBI)	Loans provided for finances in solar power projects	Active only for Rooftop PV and Small Solar Power Generation Program (RPSSGP)
2009	Feed-In Tariff (FIT)	Purchase agreement policy tool designed to promote investment in renewable energy sources	Announced on a central level in 2009 and revised in 2010
2010	Central Electricity Regulatory Commission (CERC) guidelines for Renewable Energy Certificates (RECs)	Market tool to tackle the inequality of demand and supply.	Announced in 2011 by CERC; proposed to be removed by the Ministry of Power
2011	National Clean Energy Fund (NCEF) Renewable Regulatory Fund (RRF)	Provide fund for project related to sustainable sources of energy	Power Purchase Agreement (PPA) a bonding of power distribution companies. Annual power phase cost calculation declared time to time by Central Electricity Regulatory Commission (CERC)
2013	Accelerated Depreciation	Accelerated depreciation is	GBI reinstated in

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

	(AD) Generation Based Incentive (GBI)	the method for financial accounting while generation-based incentive is a scheme of MNRE (Ministry of new and renewable energy) for wind and solar projects	2013. A payment above Feed in tariff, subject to cumulative ₹10 million/MW of installed capacity. Operational guideline for AD declared by IREDA and reinstated in 2014 with 40% growth Revised operational guideline of GBI in 2015 declared by IREDA Formation of Wind-Solar Hybrid Policy draft in 2016 by MNRE resulting in AD reduction.
2017	The Ministry of Power (MOP) recommendation	Ministry of Power proposed to remove REC Declaration of Wind-Solar Hybrid Policy by MNRE	Guidelines to power up Solar PV Power Plants.
2020	National Solar Mission (Phase 3 2017 - 2022)	To establish India as a global leader in solar energy A target of 20 GW by 2022.	34,627.82 MW installed capacity and 22 % annual growth

The solar power plant depends upon geographic location; therefore, policies vary from the state-to-state basis to take maximum advantage of available solar energy.

Each state and UT developed a vision to take maximum advantage of solar power and based on that, keep targets to generate electricity through solar energy. Hence each state and UT have its own energy policy to generate more power. Table 17 present a detailed description of the various policies related to solar power in the different states and union territories.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Table 17 State government and Union Territory solar policy progress and development [100].

State	Policy incentives	Targets	Target segment
Telangana	30% subsidy on capital for installation from MNRE and 20% subsidy on installation of RTPV up to 3kW capacity for domestic sector.	5000 MW by 2021 and 3800 MW has been installed	Aiming for decentralized solar projects. 80 % for industrial users allowed
Chhattisgarh	Capital investment subsidy of 30 % by MNRE and subsidy of 40% by the state government.	500-1000 MW by 2018 (not yet achieved)	Residential, Commercial & Industrial consumers
Gujarat	Residential rooftop capacity 80% of sanctioned load. For rooftop systems, subsidy of 40 % up to 3kW systems and 20% subsidy for systems between 3-10kW systems	8024 MW by 2022 3200 MW from rooftop installations	Small-scale distributed solar sector
Karnataka	30 % capital subsidy for non-industrial panels. Domestic consumer tariff for projects of 1kW to 10kW	6000MW by 2021 2400MW from rooftop installations (till now)	Domestic consumers
Kerala	Net metering and electricity duty exempted	500 MW by 2022 150 MW rooftop photovoltaic from domestic and agriculture consumers by 2022 and 2500 MW by 2030	Decentralized solar rooftop panel installations
Tamil Nadu	Systems can now be sized to 100% of sanctioned load (112 kW per consumer) and Consumer category solar installations: Exempted from 2 years electricity tax	9000 MW (half of peak day demand) by 2023	5400 MW - from utility-category customers 3600 MW - from consumer-category customers
Rajasthan	Electricity above 50 units - paid for by DISCOM at applicable rates for domestic customers, excess above 100 units will be purchased at INR 3.14/kWh 40% capital subsidy for system size up to 3kW and 20% subsidy for system size from 3kW up to 10kW	30000 MW by 2024-2025	Decentralized grid projects, rooftops and solar parks
West Bengal	Net metering w/tariff available for 3 phase 5kW+ systems. 50 KW solar rooftop systems.	4500MW by 2022.	All consumers
Haryana	A solar system with a capacity of 3-5% of the connected load	3200 MW by the year 2021-22	All consumers

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

Due to the strong emphasis on solar energy utilization, there is a significant increase in power generation through solar energy happens in various states and UTs.

As discussed earlier, many measures were taken to increase the use of solar power in India but, there have been delays and conflicts in implementing solar projects and other renewable projects. As the topic is limited to Solar power development in India, solar power policies in different major solar hotspot Indian states for the last 10^{th} five-year summarized in Table 17 [100]. Tables 16 and 17 reveals the results achieved still lack the primary optimistic target. The main problem here is the less knowledge of risk management. It is not evident that everything goes well in one direction only. Talented architects, contractors, skilled engineers' factors are also to be considered. From the following Table 18, it can be inferred the risks involved which led to the setback of solar assets [101]. Suggestions for risks encountered are as follows:

I. Risks encountered should be adequately studied, understood, and then renewable power projects should be developed.
II. Awareness about corruption and measures to stop it should be increased.
III. Better education facilities to increase the level of skilled human resources.
IV. Encouraging new minds to discover new and better technology by research work.
V.

Table 18 Classification of risk factor for solar asset [101].

Classification of Risks	Risks/ Barriers Description
Policy Risks	▪ Not economical ▪ Frequent government policy changes
Financial and equity risks	▪ Irregular taxation ▪ Large costs or capital investments ▪ Difficulties in getting funds from the govt.
Technology	▪ Less production of fuel ▪ Lack of knowledge ▪ Delay in power generation due to imports
Maintenance and operation	▪ There may arise many problems like defect in machines, deterioration of quality, etc.
Human resource risks	▪ Skilled labourers, engineers, etc., are not easily available.
Transparency Risk	▪ People surge for more profit and end up in benefiting themselves through corruption and India still has its roots deep into corruption.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

The detail is presented in Table 14. Among 28 states, Karnataka generates 7274.92 MW from solar power, which is the highest among all states. Punjab generates 947.10 MW from solar power, which is the lowest among all the states. The positive point of this state is that state government started giving priority in increasing the solar power capacity shortly.

Since the CSP (Concentrated Solar Power) has a very elaborate strict site requirement, very few sites are available for CSP. However, electricity can also be generated for a solar photovoltaic panel installed on the rooftop or the ground. The site requirement is significantly less; therefore, government gives strong emphasis on it. Table 15 presents the entire solar rooftop installed capacity in India [96]. The maximum solar rooftop systems are being installed in Gujrat state with a total amount of 301 MW and NCT of Delhi state has minimum installed capacity 109.80 MW [98].

There is rapid development in the generation of electricity by solar energy due to mainly various government policies. A few of different government policies introduced in India are mentioned below:

- **I.** National Electricity Policy 2015 (after revival from 2005) - enables SERCs to uphold their tariff as per their preference to generate electricity cost-effectively.
- **II.** Ministry of New and Renewable Energy (MNRE) decided the rules and regulations for renewable assets through a grid connection.
- **III.** National Solar Mission-to increases the productivity of solar power generation by discovering efficient hotspots having high potential.
- **IV.** Jawaharlal Nehru National Solar Mission (JNSSM) - has decided to increase solar thermal and solar PV (both off-grid and grid-connection).
- **V.** The State Electricity Regulatory Commissions (SERCs) and Central Electricity Regulatory Commission (CERC) - decided various guidelines for increased solar utilization and productivity.
- **VI.** National tariff policy, which includes increase in the number of certain RPOs (Renewable Purchase Obligations).

8. Future Scope

The detailed study shows massive capability of solar power establishment in India. Every year the Ministry of New and Renewable Energy (MNRE) India circulate a detailed report on the

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

solar power plant, power generation, challenges, and government policy. From the report, it can be concluded that persistent growth in the field of solar power could provide energy security soon.

The following are some of the recommendations for future study in this area: -

I. Most of the research has been carried on the technical and physical aspects to harness solar energy. In order to provide complete analysis, the economic part of solar-driven power plants need to be evaluated.
II. Solar Power generation has the potency to meet the 60-65% power requirement in India. MNRE could carry out the future of installing large projects in Rajasthan and Jammu and Kashmir. In comparison, others should identify a suitable site for the installation of a large-scale solar power plants.
III. Rooftop solar energy can be installed in the house which can save up to 50% electricity. This building energy requirement should give various solar based equipment subsidies that can further be decreased.
IV. Commercial solar panel efficiency is still low. Hence there is significant scope of improvement is required to enhance the efficiency of the solar panel.
V. The cost of installation of concentrated solar power is very high. Hence, the government should encourage and fund researchers to bring a cost-effective model that should be eased to use, cost effective, compact. The industry should also convert the prototype model to a commercial product so that this can be used effectively.
VI. The government should encourage entrepreneurs to invest their money in solar power plant and create an atmosphere of ease to do business in this area.

9. Conclusion

Among all the renewable resources, solar energy has a great scope for India to attain energy security. If the utilisation of solar energy is continued, it would be a significant part of the total electricity generation in India.

Based on the above study, the following conclusion can be drawn:

I. Energy is a basic need, and our country is not fully equipped to fulfil the necessary energy demand in the future. Thus, it is very important for India to switch over the renewable sources of energy as fossil fuels are also in limited quantities.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

II. India's geographical factors favour the use of solar energy for power generation. Considering this, many projects have started in different states under the supervision of the government of India as well as the policies decided for it.

III. Due to certain problems that persist on the inner wall, there has been a great delay in accessing the targets. Thus, people should be made aware of the risks involved and should be encouraged to use renewable energy sources as they have a great potential to develop Solar project.

IV. Solar power plants are in abundance in India, thus making it one of the top producers of renewable energy power by using solar energy.

Thus, we can conclude based on the above-mentioned reasons that there is immediate need of renewable assets.

REFERENCES

1. Kumar, M. Dinesh, Christopher A. Scott, and O. P. Singh. "Inducing the shift from flat-rate or free agricultural power to metered supply: Implications for groundwater depletion and power sector viability in India." Journal of Hydrology 2011; **409(1-2)**: 382-394.
2. Singh, Anoop. "Power sector reform in India: current issues and prospects." Energy policy 2006; **34(16)**: 2480-2490.
3. Dixit, Shantanu, Subodh Wagle, and Girish Sant. "The real challenge in power sector restructuring: instilling public control through transparency, accountability and public participation (TAP)." Energy for Sustainable Development 2001; **5(3)**: 95-102.
4. Ahmad, Asim, and Om Prakash. "Performance evaluation of a solar greenhouse dryer at different bed conditions under passive mode." Journal of Solar Energy Engineering 2020; **142(1)**: 1-10.
5. Ahmad, Asim, Om Prakash, and Anil Kumar. "Drying kinetics and economic analysis of bitter gourd flakes drying inside hybrid greenhouse dryer." Environmental Science and Pollution Research 2021: 1-15.
6. Ahmad, Asim, and Om Prakash. "Thermal analysis of north wall insulated greenhouse dryer at different bed conditions operating under natural convection mode." Environmental Progress & Sustainable Energy 2019; **38(6):** e13257.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

7. Akoijam, Amitkumar Singh, and Venni V. Krishna. "Exploring the Jawaharlal Nehru national solar mission (JNNSM): Impact on innovation ecosystem in India." African Journal of Science, Technology, Innovation and Development 2017; **9(5)**: 573-585.
8. Patra, Swapan Kumar, and Mammo Muchie, eds. Science, Technology and Innovation in BRICS Countries. Routledge, 2020.
9. Kaur, Ravneet. "An Evaluative Study of Solar Photovoltaic Energy (Off-Grid) Development Programmes in India–A Study of Punjab, Haryana and Himachal Pradesh." Journal of Economic Policy and Research 2019; **15(1)**: 42-59.
10. Bose, Debajyoti, Devender K. Saini, Monika Yadav, Saurabh Shrivastava, and Nitish Parashar. "Review of sustainable grid-independent renewable energy access in remote communities of India." Integrated Environmental Assessment and Management 2021; **17(2)**: 364-375.
11. Rathore, Pushpendra Kumar Singh, Shailendra Rathore, Rudra Pratap Singh, and Sugandha Agnihotri. "Solar power utility sector in india: Challenges and opportunities." Renewable and Sustainable Energy Reviews 2018; **81**: 2703-2713.
12. Lewis, Nathan S. "Research opportunities to advance solar energy utilization." Science mag. org. 2016; **351**: 6271.
13. Narula, Kapil, B. Sudhakara Reddy, Shonali Pachauri, and S. Mahendra Dev. "Sustainable energy security for India: An assessment of the energy supply sub-system." Energy Policy 2017; **103**: 127-144.
14. Ahmad, Asim, and Om Prakash. "Development of mathematical model for drying of crops under passive greenhouse solar dryer." Materials Today: Proceedings 2021; **47**: 6227-6230.
15. Abdulgafar, Sayran A., Omar S. Omar, and Kamil M. Yousif. "Improving the efficiency of polycrystalline solar panel via water immersion method." International Journal of Innovative Research in Science, Engineering and Technology 2014; **3(1)**: 8127-8132.
16. Belkin, Vladimir. "Cycles of Oil (1970–2017), Metals (1980–2017) Price and the Eleven–Year Solar Cycle: Statistics of Strong Relations." Chelyabinsk Humanist 2018; **1**: 42.
17. Vivoda, Vlado. "Evaluating energy security in the Asia-Pacific region: A novel methodological approach." Energy policy 2010; **38(9)**: 5258-5263.
18. Höök, Mikael, Junchen Li, Kersti Johansson, and Simon Snowden. "Growth rates of global energy systems and future outlooks." Natural Resources Research 2012; **21(1):** 23-41.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

19. Szabo, Sylvia, Dilruba Begum, Sate Ahmad, Zoe Matthews, and Peter Kim Streatfield. "Scenarios of population change in the coastal Ganges Brahmaputra delta (2011-2051)." Asia-Pacific Population Journal 2015; **30(2)**: 51-72.
20. Mittal, Shivika, Hancheng Dai, Shinichiro Fujimori, and Toshihiko Masui. "Bridging greenhouse gas emissions and renewable energy deployment target: comparative assessment of China and India." Applied Energy 2016; **166**: 301-313.
21. Tripathi, Lata, A. K. Mishra, Anil Kumar Dubey, C. B. Tripathi, and Prashant Baredar. "Renewable energy: An overview on its contribution in current energy scenario of India." Renewable and Sustainable Energy Reviews 2016; **60**: 226-233.
22. Bhide, Anjali, and Carlos Rodríguez Monroy. "Energy poverty: A special focus on energy poverty in India and renewable energy technologies." Renewable and Sustainable Energy Reviews 2011; **15(2)**: 1057-1066.
23. Kumar, Rajesh, and Arun Agarwala. "Renewable energy certificate and perform, achieve, trade mechanisms to enhance the energy security for India." Energy Policy 2013; **55**: 669-676.
24. Kiesecker, Joseph, Sharon Baruch-Mordo, Mike Heiner, Dhaval Negandhi, James Oakleaf, Christina Kennedy, and Pareexit Chauhan. "Renewable energy and land use in India: a vision to facilitate sustainable development." Sustainability 2019; **12(1):** 281.
25. Bajpai, S. and Kidwai, N.R. Renewable energy education in India. Comparative professional pedagogy 2017; **7(4):** 103-113.
26. Paramati, Sudharshan Reddy, Avik Sinha, and Eyup Dogan. "The significance of renewable energy use for economic output and environmental protection: evidence from the Next 11 developing economies." Environmental Science and Pollution Research 2017; 24(15): 13546-13560.
27. Namahoro, Jean Pierre, Qiaosheng Wu, Haijun Xiao, and Na Zhou. "The impact of renewable energy, economic and population growth on CO2 emissions in the East African region: evidence from common correlated effect means group and asymmetric analysis." Energies 2021; **14(2):** 312.
28. Molotoks, Amy, Pete Smith, and Terence P. Dawson. "Impacts of land use, population, and climate change on global food security." Food and Energy Security 2021; **10(1):** e261.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

29. Roy, Brototi, and Anke Schaffartzik. "Talk renewables, walk coal: The paradox of India's energy transition." Ecological Economics 2021; **180**: 106871.
30. Rout, Auroshis, Brijesh Mainali, Suneet Singh, Chetan Singh Solanki, and Govind S. Bhati. "Assessing the financial sustainability of rural grid electrification pathway: A case study of India." Sustainable production and consumption 2021; **25**: 27-42.
31. Singh, Sukhwinder, and Ms Gurjeet Kaur. "India's Electricity Sector: Surging Growth, Sustainability Issues and Policy Alternatives Ahead." Journal of Economics 2019; **7(4):** 72-82.
32. Godil, Danish Iqbal, Arshian Sharif, Muhammad Iftikhar Ali, Ilhan Ozturk, and Rimsha Usman. "The role of financial development, R&D expenditure, globalization and institutional quality in energy consumption in India: New evidence from the QARDL approach." Journal of environmental management 2021; **285**: 112208.
33. Garg, Vijay K., Sudhir Sharma, and Dinesh Kumar. "Design and analysis of a microgrid system for reliable rural electrification." International Transactions on Electrical Energy Systems 2021; **31(2):** e12734.
34. Ribeiro, F.G., Souza, A.P. and Carraro, A. Rural electrification and agricultural family time allocation decisions. Applied Economics 2021: 1-19.
35. Tongia, Rahul, Anurag Sehgal, and Puneet Kamboj. Future of Coal in India: Smooth Transition or Bumpy Road Ahead?. Notion Press, 2020.
36. Cropper, Maureen, Ryna Cui, Sarath Guttikunda, Nate Hultman, Puja Jawahar, Yongjoon Park, Xinlu Yao, and Xiao-Peng Song. "The mortality impacts of current and planned coal-fired power plants in India." Proceedings of the National Academy of Sciences 2021; **118(5):** 1-7.
37. Kumar, Mukul. "Disassembling Coal: Finance Capital, Environmental Law, and the Right to Information in South India." Antipode 2021; **53(4):** 1124-1142.
38. Mallick, Debarshi, Pinakeswar Mahanta, and Vijayanand S. Moholkar. "Synergistic effects in gasification of coal/biomass blends: analysis and review." Coal and biomass gasification 2018: 473-497.
39. Islam, Mohammad Maksimul, Roshan Wathore, Hisham Zerriffi, Julian D. Marshall, Rob Bailis, and Andrew P. Grieshop. "In-use emissions from biomass and LPG stoves measured

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

during a large, multi-year cookstove intervention study in rural India." Science of The Total Environment 2021; **758**: 143698.

40. Parikh, Raj, Sowmya R. Rao, Rakesh Kukde, George T. O'Connor, Archana Patel, and Patricia L. Hibberd. "Assessing the Respiratory Effects of Air Pollution from Biomass Cookstoves on Pregnant Women in Rural India." International Journal of Environmental Research and Public Health 2021; **18(1):** 183.
41. Prakash, Vrishab, Sajal Ghosh, and Kakali Kanjilal. "Costs of avoided carbon emission from thermal and renewable sources of power in India and policy implications." Energy 2020; **200**: 117522.
42. Srivastav, Asheem. Energy Dynamics and Climate Mitigation: An Indian Perspective. Springer Nature, 2020.
43. Habib, Md, and Rahat Khan. "Environmental impacts of coal-mining and coal-fired power-plant activities in a developing country with global context." In Spatial modeling and assessment of environmental contaminants, Springer, Cham 2021: 421-493.
44. Das, Sayan, Avishek Ray, and Sudipta De. "Optimum combination of renewable resources to meet local power demand in distributed generation: A case study for a remote place of India." Energy 2020; **209:** 118473.
45. Burke, Paul J., Jinnie Widnyana, Zeba Anjum, Emma Aisbett, Budy Resosudarmo, and Kenneth GH Baldwin. "Overcoming barriers to solar and wind energy adoption in two Asian giants: India and Indonesia." Energy Policy 2019; **132:** 1216-1228.
46. Kaur, Ravneet. "An Analytical Study of Policies for Promoting Energy Efficiency in India with Reference to Industrial Sector." Journal of Economic Policy and Research 2020; **15(2):** 54-70.
47. Mangla, Sachin Kumar, Sunil Luthra, Suresh Jakhar, Sumeet Gandhi, Kamalakanta Muduli, and Anil Kumar. "A step to clean energy-Sustainability in energy system management in an emerging economy context." Journal of Cleaner Production 2020; **242:** 118462.
48. Mousinho, Maria Cândida Arrais De Miranda, Ednildo Andrade Torres, Silvio Alexandre Beisl Vieira De Melo, Dinabandhu Sahoo, and Marcelo Santana Silva. "Provoking Energy Cooperation in the South Countries: Renewable Energy in Brazil and India." International Journal of Energy Economics and Policy 2017; **7(4):** 269.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

49. Yang, Zihui, and Yongliang Zhao. "Energy consumption, carbon emissions, and economic growth in India: Evidence from directed acyclic graphs." Economic Modelling 2014; **38**: 533-540.
50. Blondeel, Mathieu, and Thijs Van de Graaf. "Toward a global coal mining moratorium? A comparative analysis of coal mining policies in the USA, China, India and Australia." Climatic Change 2018; **150(1):** 89-101.
51. Elavarasan, Rajvikram Madurai, G. M. Shafiullah, Nallapaneni Manoj Kumar, and Sanjeevikumar Padmanaban. "A state-of-the-art review on the drive of renewables in Gujarat, state of India: present situation, barriers and future initiatives." Energies 2019; **13(1):** 40.
52. Murshed, Muntasir, and Muntaha Masud Tanha. "Oil price shocks and renewable energy transition: Empirical evidence from net oil-importing South Asian economies." Energy, Ecology and Environment 2021; **6(3):** 183-203.
53. Rej, Soumen, and Barnali Nag. "Energy crossroads and prioritization of energy choices: the case of India." OPEC Energy Review 2021; **45(1):** 135-158.
54. Sahoo, Manoranjan, Hrushikesh Mallick, and Mantu Kumar Mahalik. "What deteriorates India's current account balance: crude oil imports or gold imports?." Applied Economics Letters 2020; **27(13):** 1107-1111.
55. Sultan, Zafar Ahmad, Tarek Tawfek Yousef Alkhateeb, and Mahmoud Mohamed Fawaz. "Empirical investigation of relationship between oil price and inflation: The case of India." International Journal of Energy Economics and Policy 2020; **10(3):** 90.
56. Hawaldar, Iqbal Thonse, Rajesha TM, Lokesh Lokesh, and Adel Sarea. "Causal Nexus Between the Anomalies in the Crude Oil Price and Stock Market." International Journal of Energy Economics and Policy 2020; **10(3):** 233-238.
57. Kaushik, Nikhil, and Sunil Kumar. "Price, Income and Cross-elasticity of Crude Oil Demand: A VECM approach to Indian Economy." In e-journal-First Pan IIT International Management Conference–2018. 2020.
58. Swetha, Sunkar, Vani P. Elakiya, Barret K. Ammonica, Nachiyar C. Valli, and P. Prakash. "Degradation of crude oil using the indigenous isolate Bacillus sp SEA18." 2020.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

59. Chabhadiya, Karan, Rajiv Ranjan Srivastava, and Pankaj Pathak. "Growth projections against set-target of renewable energy and resultant impact on emissions reduction in India." Environmental Engineering Research 2021; **26(2):** 46-63.
60. Wang, Qiang, and Yi Liu. "India's renewable energy: New insights from multi-regional input output and structural decomposition analysis." Journal of Cleaner Production 2021; **283**: 124230.
61. Baruah, Abhinandan, Mousumi Basu, and Deeshank Amuley. "Modeling of an autonomous hybrid renewable energy system for electrification of a township: A case study for Sikkim, India." Renewable and Sustainable Energy Reviews 2021; **135**: 110158.
62. Surie, Gita. "Fostering Sustainability through Ecosystems for Renewable Energy in India." Journal of Sustainability Research 2020; **2(1):** 1-30.
63. Madurai Elavarasan, Rajvikram, Leoponraj Selvamanohar, Kannadasan Raju, Raghavendra Rajan Vijayaraghavan, Ramkumar Subburaj, Mohammad Nurunnabi, Irfan Ahmad Khan et al. "A holistic review of the present and future drivers of the renewable energy mix in Maharashtra, state of India." Sustainability 2020; **12(16):** 6596.
64. Das, Partha, Parul Mathuria, Rohit Bhakar, Jyotirmay Mathur, Amit Kanudia, and Anoop Singh. "Flexibility requirement for large-scale renewable energy integration in Indian power system: Technology, policy and modeling options." Energy Strategy Reviews 2020; **29**: 100482.
65. Shastri, Shruti, Geetilaxmi Mohapatra, and A. K. Giri. "Economic growth, renewable and nonrenewable energy consumption nexus in India: Evidences from nonlinear ARDL approach and asymmetric causality analysis." International Journal of Energy Sector Management 2020.
66. Sawhney, Aparna. "Striving towards a circular economy: climate policy and renewable energy in India." Clean Technologies and Environmental Policy 2021; **23(2):** 491-499.
67. Das, Alok, Hardik K. Jani, Garlapati Nagababu, and Surendra Singh Kachhwaha. "Wind and solar power deployment in India: economic aspects and policy implications." African Journal of Science, Technology, Innovation and Development 2021; **13(3):** 357-375.
68. Jain, Anjali, Partha Das, Sumanth Yamujala, Rohit Bhakar, and Jyotirmay Mathur. "Resource potential and variability assessment of solar and wind energy in India." Energy 2020; **211**: 118993.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

69. Deep, Sneh, Arnab Sarkar, Mayur Ghawat, and Manoj Kumar Rajak. "Estimation of the wind energy potential for coastal locations in India using the Weibull model." Renewable Energy 2020; **161**: 319-339.
70. Thapar, Sapan, and Seema Sharma. "Factors impacting wind and solar power sectors in India: A survey-based analysis." Sustainable Production and Consumption 2020; **21**: 204-215.
71. Badruddin, Irfan Anjum. "Feasibility Study of Wind-Diesel Stand-Alone System for a Small Village in India." International Journal of Integrated Engineering 2020; **12(1):** 225-232.
72. Chaurasiya, Prem Kumar, Vilas Warudkar, and Siraj Ahmed. "Wind energy development and policy in India: A review." Energy Strategy Reviews 2019; **24**: 342-357.
73. Gergan, Mabel Denzin. "Disastrous hydropower, uneven regional development, and decolonization in India's Eastern Himalayan borderlands." Political Geography 2020; **80**: 102175.
74. Mishra, Mukesh Kumar, Nilay Khare, and Alka Bani Agrawal. "Small hydro power in India: Current status and future perspectives." Renewable and sustainable energy reviews 2015; **51**: 101-115.
75. Satpathy, Preseela, and Chinmay Pradhan. "Biogas as an alternative to stubble burning in India." Biomass Conversion and Biorefinery 2020: 1-12.
76. Singh, Shri Raj Kumar, Shri Indu Shekhar Chaturvedi, Shri Aniruddha Kumar, Shri Arvind Kumar, Vimalendra Anand Patwardhan, Dinesh Dayanand Jagdale, Shri Amitesh Kumar Sinha, and Shri Bhanu Pratap Yadav. "Ministry of new and renewable energy (Mnre)." Water and Energy International 2020; **63(8):** 6-8.
77. Elavarasan, Rajvikram Madurai, G. M. Shafiullah, Sanjeevikumar Padmanaban, Nallapaneni Manoj Kumar, Annapurna Annam, Ajayragavan Manavalanagar Vetrichelvan, Lucian Mihet-Popa, and Jens Bo Holm-Nielsen. "A comprehensive review on renewable energy development, challenges, and policies of leading Indian states with an international perspective." IEEE Access 2020; **8:** 74432-74457.
78. Swain, Shri Jatindra Nath, Shri C. Kannan, Shri SK Mishra, Shri Manoj Mathur, Shri Amitesh Kumar Sinha, Shri Vimalendra Anand, Shri Mukesh Kumar Mishra, and Rashmi Singh Patwardhan. "Solar Energy Corporation of India Limited (A Government of India Enterprise) Board of Directors." Water and Energy International 2020; **63(8):** 37-41.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

79. Ramachandra, T. V., Rishabh Jain, and Gautham Krishnadas. "Hotspots of solar potential in India." Renewable and sustainable energy reviews 2011; **15(6)**: 3178-3186.
80. Irfan, Muhammad, Zhen-Yu Zhao, Muhammad Ikram, Naeem Gul Gilal, Heng Li, and Abdul Rehman. "Assessment of India's energy dynamics: Prospects of solar energy." Journal of Renewable and Sustainable Energy 2020; **12(5)**: 053701.
81. Ershad, Ahmad Murtaza, Robert Pietzcker, Falko Ueckerdt, and Gunnar Luderer. "Managing power demand from air conditioning benefits solar PV in India scenarios for 2040." Energies 2020; **13(9)**: 2223.
82. Balat, Mustafa. "Solar technological progress and use of solar energy in the world." Energy Sources, Part A 2006; **28(10)**:979-994.
83. Salam, Rashiqa Abdul, Khuram Pervez Amber, Naeem Iqbal Ratyal, Mehboob Alam, Naveed Akram, Carlos Quiterio Gómez Muñoz, and Fausto Pedro García Márquez. "An overview on energy and development of energy integration in major South Asian countries: The building sector." Energies 2020; **13(21)**: 5776.
84. Mishra, Tripti, Amit Rabha, Ujjwal Kumar, Kusum Arunachalam, and Vijay Sridhar. "Assessment of solar power potential in a hill state of India using remote sensing and Geographic Information System." Remote Sensing Applications: Society and Environment 2020; **19**: 100370.
85. Purohit, Ishan, and Pallav Purohit. "Technical and economic potential of concentrating solar thermal power generation in India." Renewable and Sustainable Energy Reviews 2017; **78**: 648-667.
86. Shukla, Akash Kumar, K. Sudhakar, Prashant Baredar, and Rizalman Mamat. "Solar PV and BIPV system: Barrier, challenges and policy recommendation in India." Renewable and Sustainable Energy Reviews 2018; **82**: 3314-3322.
87. Kumar, Nallapaneni Manoj, Ramjee Prasad Gupta, Mobi Mathew, Arunkumar Jayakumar, and Neeraj Kumar Singh. "Performance, energy loss, and degradation prediction of roof-integrated crystalline solar PV system installed in Northern India." Case Studies in Thermal Engineering 2019; **13**: 100409.
88. Thotakura, Sandhya, Sri Chandan Kondamudi, J. Francis Xavier, Ma Quanjin, Guduru Ramakrishna Reddy, Pavan Gangwar, and Sri Lakshmi Davuluri. "Operational performance

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

of megawatt-scale grid integrated rooftop solar PV system in tropical wet and dry climates of India." Case Studies in Thermal Engineering 2020;**18**:100602.

89. Malan, Anish. "Coupled optical and thermal analysis of large aperture parabolic trough solar collector." International Journal of Energy Research 2021;**45(3)**:4630-4651.
90. Fathabadi, Hassan. "Novel low-cost parabolic trough solar collector with tpct heat pipe and solar tracker: Performance and comparing with commercial flat-plate and evacuated tube solar collectors." Solar Energy 2020; **195**: 210-222.
91. Malan, Anish, and K. Ravi Kumar. "Optical Modeling of Parabolic Trough Solar Collector." In Proceedings of the 7th International Conference on Advances in Energy Research, Springer, Singapore 2021; 81-89.
92. Sisodia, Anil Kumar, and Ram Kumar Mathur. "Performance analysis of photovoltaic module by dust deposition in Western Rajasthan." Indian J Sci Technol 2020;**13(8)**: 921-933.
93. Madurai Elavarasan, Rajvikram, Leoponraj Selvamanohar, Kannadasan Raju, Raghavendra Rajan Vijayaraghavan, Ramkumar Subburaj, Mohammad Nurunnabi, Irfan Ahmad Khan et al. "A holistic review of the present and future drivers of the renewable energy mix in Maharashtra, state of India." Sustainability 2020; **12(16)**: 6596.
94. Jolly, Kuljeet, Umar Muhammad Modibbo, Jahangir Chauhan, and Mohd Ansari. "A Cost-Based Optimization Modelling of Solar Power Generation in India for Sustainable Development." In Computational Modelling in Industry 4.0, Springer, Singapore 2022: 35-47.
95. KhareSaxena, Anupama, Seema Saxena, and K. Sudhakar. "Solar energy policy of India: An overview." CSEE Journal of Power and Energy Systems 2020.
96. Gupta, Sandeep, and Aamir Khan Nurkhani. "Financial Analysis of Solar Energy Development in India: Potential, Challenges and Policies." In Computational Network Application Tools for Performance Management, Springer, Singapore 2020: 259-267.
97. Bandi, Venkata, Tiia Sahrakorpi, Jukka Paatero, and Risto Lahdelma. "Touching the invisible: Exploring the nexus of energy access, entrepreneurship, and solar homes systems in India." Energy Research & Social Science 2020; **69:** 101767.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

98. Bhatt, Brijesh, and Anoop Singh. "Stakeholders' role in distribution loss reduction technology adoption in the Indian electricity sector: An actor-oriented approach." Energy Policy 2020; **137:** 111064.
99. Sabaawi, Ahmed MA, Charalampos C. Tsimenidis, and Bayan S. Sharif. "Analysis and modeling of infrared solar rectennas." IEEE Journal of Selected Topics in Quantum Electronics 2013; **19(3)**: 9000208-9000208.
100. Sharma, Sumit, and Yog Raj Sood. "Microgrids: A review of status, technologies, software tools, and issues in Indian power market." IETE Technical Review 2020:1-22.
101. Liu, X., & Zeng, M. Renewable energy investment risk evaluation model based on system dynamics. Renewable and Sustainable Energy Reviews 2017;**73:**782-788.

ORIGINAL UNEDITED MANUSCRIPT

Downloaded from https://academic.oup.com/ijlct/advance-article/doi/10.1093/ijlct/ctac075/6649853 by Bangalore University user on 04 August 2022

arXiv:2208.00753v1 [math.CV] 1 Aug 2022

THEORY OF CERTAIN NON-UNIVALENT ANALYTIC FUNCTIONS

KAMALJEET GANGANIA

Abstract. We investigate the non-univalent function's properties reminiscent of the theory of univalent starlike functions. Let the analytic function $\psi(z)=\sum_{i=1}^{\infty}A_iz^i$, $A_1\neq 0$ be univalent in the unit disk. Non-univalent functions may be found in the class $\mathcal{F}(\psi)$ of analytic functions f of the form $f(z)=z+\sum_{k=2}^{\infty}a_kz^k$ satisfying $(zf'(z)/f(z)-1)\prec\psi(z)$. Such functions, like the Ma and Minda classes of starlike functions, also have nice geometric properties. For these functions, growth and distortion theorems have been established. Further, we obtain bounds for some sharp coefficient functionals and establish the Bohr and Rogosinki phenomenon for the class $\mathcal{F}(\psi)$. Non-analytic functions that share properties of analytic functions are known as Poly-analytic functions. Moreover, we compute Bohr and Rogosinski's radius for Poly-analytic functions with analytic counterparts in the class $\mathcal{F}(\psi)$ or classes of Ma-Minda starlike and convex functions.

1. Introduction

Let $\mathcal{A}$ be the class of analytic functions of the form $f(z)=z+\sum_{k=2}^{\infty}a_kz^k$ in the open unit disk $\Delta:=\{z:|z|<1\}$. Let $f(z)=w$ and Γ_w be the image of Γ_z (the circle $C_r:z=re^{i\theta}$) under the function f in $\mathcal{A}$. The curve Γ_w is said to be starlike with respect to $w_0=0$ if $\arg(w-w_0)$ is a non-decreasing function of θ, that is,

$$\frac{d}{d\theta}\arg(w-w_0)\geq 0,\quad \theta\in[0,2\pi],$$

which is equivalent to

$$\frac{d}{d\theta}\arg(w-w_0)=\Re\left(\frac{zf'(z)}{f(z)}\right)\geq 0. \tag{1.1}$$

From (1.1), we note the importance of the Carathéodory functions when (1.1) is viewed in terms of subordination as:

$$\frac{zf'(z)}{f(z)}\prec\frac{1+z}{1-z}\quad(z\in\Delta), \tag{1.2}$$

where the symbol $\prec$ means for the usual subordination. In view of (1.2), let Ψ be an analytic function in Δ with positive real part, $\Psi(\Delta)$ symmetric about the real axis with $\Psi'(0)>0$ and $\Psi(0)=1$. Generalizing (1.2), Ma and Minda [28] considered the unified approach by defining the class of starlike functions

$$\mathcal{S}^*(\Psi):=\left\{f\in\mathcal{A}:\frac{zf'(z)}{f(z)}\prec\Psi(z)\right\}. \tag{1.3}$$

If the inequality (1.1) holds for each circle $|z|=r<1$, then it characterizes a special class $\mathcal{S}^*$, the class of starlike functions in Δ. It is obvious from (1) that for each $0<r<1$, the curve Γ_w is not allowed to have a loop which ensure the univalency of the function. But if for some $0\neq z\in\Delta$,

2010 *Mathematics Subject Classification.* Primary 30C45, 30C35, 30C50; Secondary 35Q30.

Key words and phrases. Poly-analytic, Bi-analytic, Non-Univalent, Starlike, Bohr and Rogosinski's inequality, Coefficient problems.

The work of Kamajeet Gangania is supported by University Grant Commission, New-Delhi, India under UGC-Ref. No.:1051/(CSIR-UGC NET JUNE 2017).

$\Re(zf'(z)/f(z)) < 0$, then f is not starlike with respect to 0, or equivalently we can say that the image curve $\Gamma_w : f(|z| = r)$ is definitely not starlike, but still it may or may not be univalent.

Therefore, view of (1.1), (1.2) and the class (1.3), if we take

$$\frac{zg'(z)}{g(z)} - 1 \prec \psi(z), \tag{1.4}$$

where ψ be the analytic univalent function in Δ such that $\psi(0) = 0$, $\psi(\Delta)$ is starlike with respect to 0. Then g satisfying (1.4) may not be univalent. In view of (1.4), Kargar et al. [16] studied a class reminiscent to the Ma-Minda class.

$$\mathcal{BS}(\beta) := \left\{ f \in \mathcal{A} : \frac{zf'(z)}{f(z)} - 1 \prec \frac{z}{1-\beta z^2},\ \beta \in [0,1) \right\},$$

where $z/(1-\beta z^2) =: \psi(z)$ (Booth lemniscate function [30] and [31]) is an analytic univalent function and symmetric with respect to the real and imaginary axes. Also, see the family of generalized pascal snails [15]. Masih et al. [29] considered the following class with $\beta \in [0, 1/2]$:

$$\mathcal{S}_{cs}(\beta) := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \frac{z}{(1-z)(1+\beta z)},\ \beta \in [0,1) \right\}.$$

Kumar and Gangania [20] studied the class

$$\mathcal{S}_{\gamma}(\eta) := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \frac{\gamma z}{(1+\eta z)^2},\ \eta \in [0,1),\ \gamma > 0 \right\}.$$

Motivated by the above classes and observations, Kumar and Gangania [20] introduced the following class:

Definition 1. [20] Let ψ be the analytic univalent function in Δ such that $\psi(0) = 0$, $\psi(\Delta)$ is starlike with respect to 0. Then

$$\mathcal{F}(\psi) := \left\{ f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1 \right) \prec \psi(z),\ \psi(0) = 0 \right\}.$$

The following function plays the role of extremal

$$f_0(z) = z \exp \int_0^z \frac{\psi(t)}{t} dt. \tag{1.5}$$

Notice that for $1 + \psi(z) \not\prec (1+z)/(1-z)$, the functions in the class $\mathcal{F}(\psi)$ may not be univalent in Δ which also implies $\mathcal{F}(\psi) \not\subseteq \mathcal{S}^*$. Authors in [20] noticed that such non-univalent analytic functions also exhibit nice geometric properties similar to class $\mathcal{S}^*(\Psi)$ such as

Theorem 1.1. [20]*(Covering theorem) Let $\min_{|z|=r} \Re\psi(z) = \psi(-r)$. If $f \in \mathcal{F}(\psi)$ and f_0 as defined in* (1.5)*, then either f is a rotation of f_0 or*

$$\{w \in \mathbb{C} : |w| \leq -f_0(-1)\} \subset f(\Delta),$$

where $-f_0(-1) = \lim_{r\to 1}(-f_0(-r))$.

Remark 1. *The class $\mathcal{F}(\psi)$ gives a new insight and a pool of opportunities to study non-univalent functions. Several problems, namely Area estimates, Logarithmic coefficient bounds, sharp coefficient bounds, hankel determinant, Zalcman functional, and several radius problems like majorization, radius of convexity, Bohr-phenomenon, differential subordination implication, preservation under convolution and some famous integral etc.* are yet to be known.

Moreover, inclusion relation like $\mathcal{S}^*(\Psi) \subseteq \mathcal{S}^*$ is not available for the class $\mathcal{F}(\psi)$. This fact makes the above discussed problems non-trivial in $\mathcal{F}(\psi)$. In fact, several implications which holds for $\mathcal{S}^*(\Psi)$ are not true for the class $\mathcal{F}(\psi)$. For instance, if $f \in \mathcal{S}^*(\Psi)$ then $\Re(f(z)/z) > 0$ for all $z \in \Delta$, see [28]. But $\Re(f(z)/z) > 0$ may not be true for $f \in \mathcal{F}(\psi)$. See Figure 1. The function

$$\hat{f}(z) = z\left(\frac{1+z\sqrt{\beta}}{1-z\sqrt{\beta}}\right)^{\frac{1}{2\sqrt{\beta}}} \in \mathcal{BS}(\beta) \Rightarrow \Re\left(\frac{\hat{f}(z)}{z}\right) > 0 \quad \text{for all } z \in \Delta.$$

However, for all $z \in \Delta$, the function

$$\kappa(z) := z\exp\left(\frac{z}{(1+\eta z)^2}\right) \in \mathcal{S}_1(\eta) \Rightarrow \Re\left(\frac{\kappa(z)}{z}\right) \not> 0, \quad \eta > 2-\sqrt{3}.$$

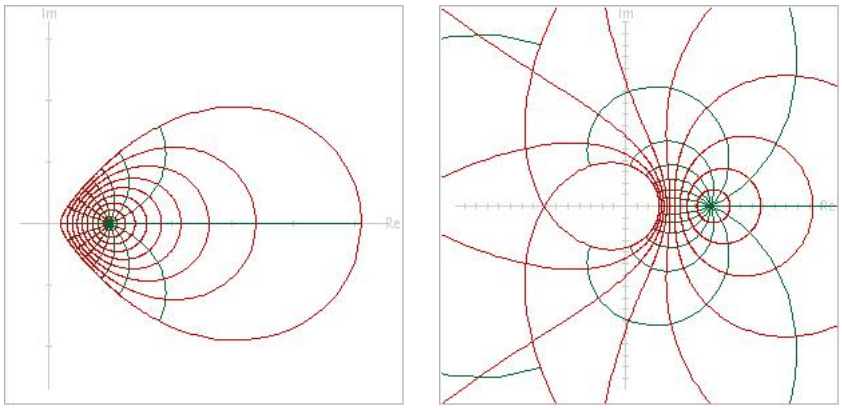


FIGURE 1. Left side graph is for $\hat{f}(z)/z$ ($\beta = 0.9$) and Right side graph is for $\kappa(z)/z$ ($\eta = 0.9$).

Likewise, Non-analytic functions with significant structure and properties similar to those satisfied by analytic functions exist as well. Non-analytic functions with such a nice structure are referred to as poly-analytic functions, which we shall discuss briefly soon in Section 2. For there interesting historical viewpoint and the application's side, we refer to read [1, Page 170]

The version of Landau's theorem for poly-analytic functions is given in

Theorem 1.2. [1, Theorem 1] *(Landau's theorem) Let $F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)$ be a poly-analytic function of order α, $\alpha \geq 2$ on Δ, where f_k are analytic such that $f_k(0) = 0$, $f_k'(0) = 1$ and $|f_k(z)| \leq M$, for all k, with $M > 1$. Then there is a constant $0 < \rho_1 < 1$ so that F is univalent in $|z| < \rho_1$. In particular, ρ_1 satisfies*

$$1 - M\left(\frac{\rho_1(2-\rho_1)}{(1-\rho_1)^2} + \sum_{k=1}^{\alpha-1} \frac{\rho_1^k(1+k+k\rho_1)}{(1-k\rho_1)^2}\right) = 0$$

and $F(\Delta_{\rho_1})$ contains a disk Δ_{R_1}, where

$$R_1 = \rho_1 - {\rho_1}^2\left(\frac{1-{\rho_1}^{\alpha-1}}{1-\rho_1}\right) - M\sum_{k=0}^{\alpha-1} \frac{{\rho_1}^{k+2}}{1-\rho_1}.$$

For other relevant work in this direction, we refer to see [1, 9, 26] and references therein.

For poly-analytic functions of order α given by $F(z)=\sum_{k=0}^{\alpha-1}\bar{z}^k f_k(z)$, where $f_k(z)=\sum_{n=0}^{\infty}a_{n,k}z^n$, Abdulhadi and El Hajj [1] defined the majorant series $M(F,r)$ by

$$M(F,r)=\sum_{n=0}^{\infty}\sum_{k=0}^{\alpha-1}|a_{n,k}|r^{k+n},$$

and proved the following Bohr's type inequality

Theorem 1.3. [1, Theorem 2] *Let $F(z)=\sum_{k=0}^{\alpha-1}\bar{z}^k f_k(z)$ be a poly-analytic function of order α, where f_k are analytic mappings for $k=0,1\cdots,\alpha-1$, such that $g_k=f_0+\bar{f_k}$ is sense-preserving in Δ and $f_k(0)=0$ for each $k=0,\cdots,\alpha-1$. Suppose that f_0 is univalent and normalized by $f_0(0)=0$, $f_0'(0)=1$. Then*

$$M(F,r)<1 \quad if \quad |z|<r_0,$$

where $r_0\approx 0.318$ is the minimum positive root of the polynomial

$$(1-r)^3-r+r^{\alpha+1}=0.$$

The classical Bohr's inequality [6] says that if $f(z)=\sum_{n=0}^{\infty}a_nz^n$ is analytic in Δ and $|f(z)|<1$ for all $z\in\Delta$, then $\sum_{n=0}^{\infty}|a_nz^n|<1$ for all z in Δ with $|z|<1/3$. The constant $1/3$ can not be improved and is known as Bohr's radius. In analogy with Bohr's inequality, there is also the notion of Rogosinski radius, however a little is known about Rogosinski radius as compared to Bohr radius, which states that, see [22, 33, 35]: if $g(z)=\sum_{k=0}^{\infty}b_k$ with $|f(z)|<1$, then for every $N\geq 1$ we have $|\sum_{k=0}^{N-1}b_kz^k|\leq 1$ for all $|z|\leq 1/2$. The radius $1/2$ is called the Rogosinski radius.

Kayumov et al. [17] studied a new quantity, called Bohr-Rogosinski sum and is given below:

$$|g(z)|+\sum_{k=N}^{\infty}|b_k||z|^k, \quad |z|=r.$$

For the case $N=1$, note that this sum is similar to the Bohr's sum, where $g(0)$ is replaced by $|g(z)|$. We also refer the readers to see [2, 5]. Now we say the family $S(f)$ has Bohr-Rogosinski phenomenon, if there exists $r_N^f\in(0,1]$ such that the inequality:

$$|g(z)|+\sum_{k=N}^{\infty}|b_k||z|^k\leq|f(0)|+d(f(0),\partial\Omega)$$

holds for $|z|=r\leq r_N^f$, where $d(f(0),\partial\Omega)$ denotes the Euclidean distance between $f(0)$ and the boundary of $\Omega=f(\Delta)$. The largest such r_N^f is called the Bohr-Rogosinski radius.

In this investigation, we obtain sharp initial coefficient's estimates along with some well known coefficient functionals for the class $\mathcal{F}(\psi)$. We also derive Bohr and Bohr-Rogosinski radius for $\mathcal{F}(\psi)$. Further, we explore the classes $\mathcal{F}(\psi)$ and $\mathcal{S}^*(\Psi)$ in connection with the poly-analytic functions to obtain Bohr and Rogosinski's type radius.

2. Radius problems: The Bohr and Rogosinski phenomenon

For our next result, we need to recall the following result of Ruscheweyh and Stankiewicz [34]:

Lemma 2.1 ([34])**.** *Let the analytic functions F and G be convex univalent in Δ. If $f\prec F$ and $g\prec G$, then*

$$f*g\prec F*G \quad (z\in\Delta).$$

We also refer to see [28] for a parallel result concerning the class $\mathcal{S}^*(\Psi)$.

Lemma 2.2. *Let $\psi(\Delta)$ be a convex domain . If $f\in\mathcal{F}(\psi)$. Then $f(z)/z\prec f_\psi(z)/z$, where $f_\psi(z)=z\exp\int_0^z(\psi(t)/t)dt$.*

Proof. Let $f \in \mathcal{F}(\psi)$. Then

$$\phi(z) := \frac{zf'(z)}{f(z)} - 1 \prec \psi(z). \tag{2.1}$$

Let us take the univalent convex function

$$g(z) = \log\left(\frac{1}{1-z}\right) = \sum_{n=1}^{\infty} \frac{z^n}{n}.$$

Thus for $f \in \mathcal{A}$, we get

$$\phi(z) * g(z) = \int_0^z \frac{\phi(t)}{t} dt \quad \text{and} \quad \psi(z) * g(z) = \int_0^z \frac{\psi(t)}{t} dt. \tag{2.2}$$

Since ψ is convex. Therefore, applying Lemma 2.1 in (2.1), we get

$$\phi(z) * g(z) \prec \psi(z) * g(z). \tag{2.3}$$

Now from (2.2) and (2.3), we obtain

$$\int_0^z \frac{\phi(t)}{t} dt \prec \int_0^z \frac{\psi(t)}{t} dt,$$

which implies that

$$\frac{f(z)}{z} := \exp \int_0^z \frac{\phi(t)}{t} dt \prec \exp \int_0^z \frac{\psi(t)}{t} dt =: \frac{f_\psi(z)}{z}.$$

This completes the proof. □

Remark 2. (Behaviour of the functions $f(z)/z$) It is worth to mention that the extremal function $f_\psi(z)/z$ may not be univalent, Carathéodory and convex in Lemma 2.2, see [20] and Figure 1. This restrict us to examine the parallel results as in the Ma and Minda classes. But still this information is useful to settle the Bohr-phenomenon for the class $\mathcal{F}(\psi)$.

To discuss the Bohr phenomenon for the class $\mathcal{F}(\psi)$, we recall the following which is the analogue of the growth theorem [28] for $\mathcal{S}^*(\Psi)$:

Lemma 2.3. [20] *(Growth Theorem) Let us suppose that* $\max_{|z|=r} \Re\psi(z) = \psi(r)$ *and* $\min_{|z|=r} \Re\psi(z) = \psi(-r)$. *Then* $f \in \mathcal{F}(\psi)$ *satisfies the sharp inequalities:*

$$r \exp\left(\int_0^r \frac{\psi(-t)}{t} dt\right) \leq |f(z)| \leq r \exp\left(\int_0^r \frac{\psi(t)}{t} dt\right), \quad (|z| = r). \tag{2.4}$$

Without loss of generality, we can have in Lemma 2.3 that $\max_{|z|=r} \Re\psi(z) = M(r)$ and $\min_{|z|=r} \Re\psi(z) = m(r)$, where M and N are functions of r for each z in Δ such that $|z| = r$.

Remark 3. In case of carathéodory functions p and q, we see that $p \prec q$, where $q(z) = (1+z)/(1-z)$ is the best dominant. But in case of functions in $\mathcal{S}$, we have no such best dominant. Therefore, it is worthy to obtain upper bound for $|f(z)|$ in general for the class $\mathcal{F}(\psi)$, where $\psi \in \mathcal{S}$. Note that

$$\max_{|z|=r} \left|\frac{zf'(z)}{f(z)}\right| \leq \frac{1-r+r^2}{(1-r)^2}.$$

This yields

$$|f'(z)| \leq \frac{(1-r+r^2)}{(1-r)^2} \exp\left(\frac{r}{1-r}\right) \quad \text{and} \quad |f(z)| \leq r \exp\left(\frac{r}{1-r}\right).$$

The equality cases hold for the non-univalent function

$$f(z) = z \exp\left(\frac{z}{1-z}\right).$$

At this conjunction in view of the remark 3, we propose the following:

***Question* 1.** Let $m(r) := \min_{|z|=r} |f(z)|$, where $f \in \mathcal{F}(\psi)$ and $\psi \in \mathcal{S}$. Find the value of $m(r)$.

In the following result, it is important to note that bounds for $|b_n|$ are not available. In rest of this section, we assume that coefficient A_i of ψ are positive.

Remark 4. In case, when all the coefficients A_i of function ψ are not positive, then without loss of generality we can appropriately replace the function $f_\psi(z) = z \exp \int_0^z (\psi(t)/t)dt := z + \sum_{n=2}^{\infty} a_n z^n$ by the analytic function $\hat{f}_\psi(z) := z + \sum_{n=2}^{\infty} |a_n| z^n$, throughout this paper.

Theorem 2.1. *Let $\psi(\Delta)$ be convex. The class $S(f)$, where $f \in \mathcal{F}(\psi)$ satisfies the Bohr-phenomenon*

$$\sum_{n=1}^{\infty} |b_n||z|^n \leq d(0, \partial f(\Delta))$$

in $|z| = r \leq \min\{1/3, r_0\}$, where r_0 is least positive root of the equation

$$f_\psi(r) - \exp \int_0^1 \frac{\psi(-t)}{t} dt = 0. \tag{2.5}$$

The result is sharp in case of $r_0 \leq 1/3$.

Proof. Since $g \in S(\mathcal{F}(\psi))$, we have $g \prec f$ for a fixed $f \in \mathcal{F}(\psi)$. From Lemma 2.3, we obtain the Koebe-radius

$$r_* = \exp \int_0^1 \frac{\psi(-t)}{t} dt$$

such that $r_* \leq d(0, \partial\Omega) = |f(z)|$ for $|z| = 1$. Also using Lemma 2.2, we have

$$\frac{f(z)}{z} \prec \frac{f_\psi(z)}{z}. \tag{2.6}$$

Recall the result [11, Lemma 2.1], which reads as: let f and g be analytic in Δ with $g \prec f$, where $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and $g(z) = \sum_{k=0}^{\infty} b_k z^k$. Then $\sum_{k=0}^{\infty} |b_k| r^k \leq \sum_{n=0}^{\infty} |a_n| r^n$ for $|z| = r \leq 1/3$. Now using this result for $g \prec f$ and (2.6), we have

$$\sum_{k=1}^{\infty} |b_k| r^k \leq r + \sum_{n=2}^{\infty} |a_n| r^n \leq f_\psi(r) \quad \text{for } |z| = r \leq 1/3.$$

Finally, to establish the inequality $\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega)$, it is enough to show $f_\psi(r) \leq r_*$. But this holds whenever $r \leq r_0$, where r_0 is the least positive root of the equation $f_\psi(r) = r_*$. Now let

$$G(r) := f_\psi(r) - r_*.$$

Note that $G'(r) = f'_\psi(r) > 0$ whenever the Taylor coefficient's of f_ψ are positive, yields unique positive root of $G(r) = 0$. Also, G is a continuous function of r in $(0, 1)$ such that

$$G(0) = f_\psi(0) - \exp \int_0^1 \frac{\psi(-1)}{t} dt < 0$$

and

$$G(1) = f_\psi(1) - \exp \int_0^1 \frac{\psi(-1)}{t} dt > 0$$

Thus the existence of the root r_0 is ensured by the Intermediate Value theorem for the continuous functions. Now if $r_0 \leq 1/3$, then for the choices of functions $g = f = f_\psi$, equality

$$\sum_{k=1}^{\infty} |b_k| r^k \leq d(f(0), \partial\Omega),$$

holds; hence this prove the sharpness part. □

Now we discuss a generalization of Bohr-Rogosinski sum for the class $\mathcal{F}(\psi)$. Also see [14, 18]. We first recall

Lemma 2.4. [14] *Let $f(z)=\sum_{n=0}^{\infty}a_nz^n$ and $g(z)=\sum_{n=0}^{\infty}b_nz^n$ be analytic in Δ. Let $\{\nu_k(r)\}_{k=0}^{\infty}$ be a sequence of non-negative functions, continuous in $[0,1)$ such that the series*

$$\sum_{n=0}^{\infty}|b_n|\nu_k(r)$$

converges locally uniformly with respect to $r\in[0,1)$. If $g\prec f$, then

$$\sum_{n=0}^{\infty}|a_n|\nu_n(r)\leq\sum_{n=0}^{\infty}|b_n|\nu_n(r)$$

for all $|z|=r\leq\frac{1}{3}$.

Theorem 2.2 (Generalized Bohr-Rogosinski sum)**.** *Let $\psi(\Delta)$ be convex. Let $\{\nu_n(r)\}_{n=1}^{\infty}$ be a non-negative sequence of continuous functions in $[0,1]$ such that*

$$\nu_1(r)+\sum_{n=2}^{\infty}\left|\frac{f_{\psi}^{(n)}(0)}{n!}\right|\nu_n(r)$$

converges locally uniformly with respect to each $r\in[0,1)$. If

$$|f(z^m)|+\nu_1(r)+\sum_{n=2}^{\infty}\left|\frac{f_{\psi}^{(n)}(0)}{n!}\right|\nu_n(r)<\exp\int_0^1\frac{\psi(-t)}{t}dt$$

and $f(z)=z+\sum_{n=2}^{\infty}a_nz^n\in\mathcal{F}(\psi)$. Then

$$|f(z^m)|+\sum_{n=1}^{\infty}|a_n|\nu_n(r)\leq d(0,\partial\Omega)\tag{2.7}$$

holds for $|z|=r\leq r_b=\min\{1/3,r_0\}$, where $m\in\mathbb{N}$, $\Omega=f(\Delta)$ and r_0 is the smallest positive root of the equation:

$$f_{\psi}(r^m)+\sum_{n=2}^{\infty}\left|\frac{f_{\psi}^{(n)}(0)}{n!}\right|\nu_n(r)=\exp\int_0^1\frac{\psi(-t)}{t}dt-\nu_1(r),$$

where

$$f_{\psi}(z)=z\exp\int_0^z\frac{\psi(t)}{t}dt.$$

Moreover, the inequality (2.7) *also holds for the class $S(f)$ in $|z|=r\leq r_b$. When $r_b=r_0$, then the radius is best possible.*

Proof. Since $f\in\mathcal{F}(\psi)$, it follows from Lemma 2.2 that

$$\frac{f(z)}{z}\prec\frac{f_{\Psi}(z)}{z}.$$

Applying Lemma 2.4, we see that

$$\sum_{n=1}^{\infty}|a_n|\nu_n(r)\leq\nu_1(r)+\sum_{n=2}^{\infty}\left|\frac{f_{\Psi}^{(n)}(0)}{n!}\right|\nu_n(r)\quad\text{for}\quad r\leq\frac{1}{3}.\tag{2.8}$$

Combining now the equation (2.8) with the Lemma 2.3 gives

$$|f(z^m)| + \sum_{n=1}^{\infty} |a_n|\nu_n(r) \leq f_\Psi(r^m) + \nu_1(r) + \sum_{n=2}^{\infty} \left|\frac{f_\Psi^{(n)}(0)}{n!}\right| \nu_n(r)$$

for $r \leq 1/3$. Hence, using the hypothesis

$$|f(z^m)| + \nu_1(r) + \sum_{n=2}^{\infty} \left|\frac{f_\psi^{(n)}(0)}{n!}\right| \nu_n(r) < \exp \int_0^1 \frac{\psi(-t)}{t} dt,$$

the inequality (2.7) holds in $|z| = r \leq \min\{r_0, 1/3\}$, where r_0 is the minimum positive root of the equation

$$f_\psi(r^m) + \sum_{n=2}^{\infty} \left|\frac{f_\psi^{(n)}(0)}{n!}\right| \nu_n(r) = \exp \int_0^1 \frac{\psi(-t)}{t} dt - \nu_1(r),$$

In case $r_0 \leq 1/3$, then for the function $f(z) = f_\Psi(z) = z + \sum_{n=2}^{\infty} \left|\frac{f_\Psi^{(n)}(0)}{n!}\right| z^n$, we have

$$d(0, \partial\Omega) = \exp \int_0^1 \frac{\psi(-t)}{t} dt$$

such that the equality

$$|f_\Psi(r_0^m)| + \sum_{n=1}^{\infty} |a_n|\nu_n(r_0) = \exp \int_0^1 \frac{\psi(-t)}{t} dt = d(0, \partial\Omega)$$

holds, that is, the radius is best possible. □

In the following, taking $m \to \infty$ and $N \to 1$ gives Bohr phenomenon for the class $\mathcal{F}(\psi)$.

Corollary 2.2.1. *Let $\psi(\Delta)$ be convex. If $f(z) = z + \sum_{n=2}^{\infty} a_n z^n \in \mathcal{F}(\psi)$. Then*

$$|f(z^m)| + \sum_{n=N}^{\infty} |a_n||z|^n \leq d(0, \partial\Omega) \tag{2.9}$$

holds for $|z| = r \leq r_b = \min\{1/3, r_0\}$, where $m \in \mathbb{N}$, $\Omega = f(\Delta)$ and r_0 is the smallest positive root of the equation:

$$f_\psi(r^m) + \sum_{n=N}^{\infty} \left|\frac{f_\psi^{(n)}(0)}{n!}\right| r^n = \exp \int_0^1 \frac{\psi(-t)}{t} dt,$$

where

$$f_\psi(z) = z \exp \int_0^z \frac{\psi(t)}{t} dt.$$

Moreover, the inequality (2.9) *also holds for the class $S(f)$ in $|z| = r \leq r_b$. When $r_b = r_0$, then the radius is best possible.*

Here below, Corollary 2.2.1 yields the Bohr-Rogosinki radius for the class $\mathcal{BS}(\beta)$.

Corollary 2.2.2. *Let $0 < \beta \leq 3 - 2\sqrt{2}$. If $f(z) = z + \sum_{n=2}^{\infty} a_n z^n \in \mathcal{BS}(\beta)$. Then*

$$|f(z^m)| + \sum_{n=N}^{\infty} |a_n||z|^n \leq d(0, \partial\Omega) \tag{2.10}$$

holds for $|z| = r \leq r_b = \min\{r_N^m(\beta), 1/3\}$, *where* $m \in \mathbb{N}$, $\Omega = f(\Delta)$ *and* $r_N^m(\beta)$ *is the smallest positive root of the equation:*

$$\hat{f}_\psi(r^m) - S_{N-1}(\hat{f}) + \hat{f}(r) = \left(\frac{1-\sqrt{\beta}}{1+\sqrt{\beta}}\right)^{\frac{1}{2\sqrt{\beta}}},$$

where

$$\hat{f}(z) = z\left(\frac{1+z\sqrt{\beta}}{1-z\sqrt{\beta}}\right)^{\frac{1}{2\sqrt{\beta}}}$$

and $S_{N-1}(\hat{f})$ *denotes* $(N-1)^{th}$ *partial sum of* $\hat{f}$ *and* $S_0 = 0$. *Moreover, the inequality* (2.10) *also holds for the class* $S(f)$ *in* $|z| = r \leq r_b$. *The radius* $r_N^m(\beta)$ *is sharp with* $m = 1$ *for all* $N \in \mathbb{N}$. *See, Figure 2*

Remark 5. Taking $m \to \infty$ and $N \to 1$ in Corollary 2.2.2 gives [20, Theorem 4].

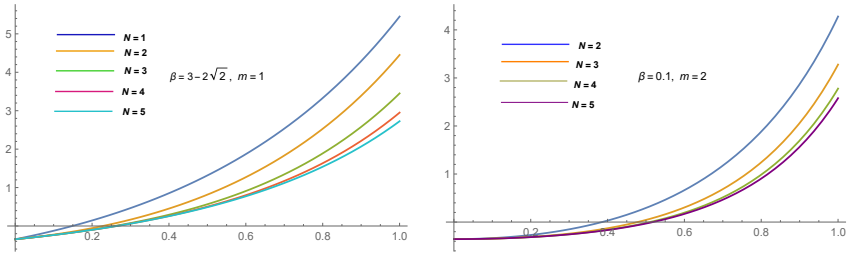


FIGURE 2. Existence of roots for different values of β, m and N in Corollary 2.2.2.

We now study poly-analytic functions in connection with the class $\mathcal{F}(\psi)$. A continuous complex-valued function F defined in a domain $D \subseteq \mathbb{C}$ is poly-analytic of order α if it obeys the generalized Cauchy-Riemann equations $\partial^\alpha F/\partial\bar{z}^\alpha = 0$. An iteration argument shows that F has the form

$$F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z),$$

where f_k are analytic functions for $k = 0, \cdots, \alpha - 1$. Throughout, it is assumed that $\alpha \geq 2$. In particular, a continuous complex-valued function $F = u + iv$ in a domain $D \subseteq \mathbb{C}$ is said to be bi-analytic if

$$\frac{\partial}{\partial\bar{z}}\left(\frac{\partial F}{\partial\bar{z}}\right) = 0.$$

Then F is a poly-analytic function of order 2. Note that $F_{\bar{z}}$ is analytic in D. In any simply connected domain D, it can be shown that F has the form

$$F(z) = \bar{z}f_1(z) + f_0(z),$$

where f_1 and f_0 are analytic functions (and obviously $f_1 = F_{\bar{z}}$). Poly-analytic functions can have properties that differ from analytic functions, such as the ability to vanish on closed curves without vanishing identically. The function $F(z) = 1 - z\bar{z}$ is one such example.

We obtain Theorem 1.3 for poly-analytic function with analytic counterparts from the classes $\mathcal{F}(\psi)$, $\mathcal{S}^*(\Psi)$ and $\mathcal{C}(\Psi)$ for the Rogosinski's sum.

Theorem 2.3. *Let* $F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)$ *be a poly-analytic function of order* α, *where* f_k *are analytic such that* $g_k = f_0 + \overline{f_k}$ *is sense-preserving in* Δ *and* $f_k(0) = 0$ *for each* $k = 0, \cdots, \alpha - 1$. *Suppose that*

(*i*) $L(r) = B_N^m(f_\psi, r)$ *if* $f_0 \in \mathcal{F}(\psi)$ *and* $\psi(\Delta)$ *is convex*
(*ii*) $L(r) = B_N^m(f_\Psi, r)$ *if* $f_0 \in \mathcal{S}^*(\Psi)$
(*iii*) $L(r) = B_N^m(g_\Psi, r)$, *where* $zg'_\Psi(z) = f_\Psi(z)$ *if* $f_0 \in \mathcal{C}(\Psi)$.

Then

$$B_N^m(F, r) = \sum_{n=N}^{m} \sum_{k=0}^{\alpha-1} |a_{n,k}| r^{k+n} < 1, \quad \text{if} \quad |z| < \min\{1/3, r_0\}$$

where $N \in \mathbb{N}$ *and* r_0 *is the minimal positive root of the equation*

$$L(r)(1 - r^\alpha) + r - 1 = 0.$$

Proof. First see that

$$B_N^m(F, r) = B_N^m\left(\sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)\right) \leq \sum_{k=0}^{\alpha-1} r^k B_N^m(f_k). \tag{2.11}$$

Since $g_k = f_0 + \overline{f_k}$ is sense-preserving in Δ for each k, it follows that

$$f'_k(z) = \omega_k(z) f'_0(z)$$

where ω_k is analytic in Δ and $|\omega_k(z)| < 1$ for all $z \in \Delta$, which further using [11, Corollary 2.1] gives that for $|z| \leq 1/3$,

$$\begin{aligned} B_N^m(f_k) &= \int_0^r B_N^m(f'_k) ds = \int_0^r B_N^m(\omega_k f'_0) ds \\ &\leq \int_0^r B_N^m(f'_0) ds = B_N^m(f_0). \end{aligned}$$

Thus, from (2.11)

$$B_N^m(F, r) \leq B_N^m(f_0) \left(\frac{1 - r^\alpha}{1 - r}\right) \quad \text{for} \quad r \leq \frac{1}{3}.$$

Now to complete the proof, we consider the following three cases:

(i) Let $f_0 \in \mathcal{S}^*(\psi)$. Then $f_0(z)/z \prec f_\psi(z)/z$ holds, which with [11, Lemma 2.1] gives

$$B_N^m(f_0) \leq B_N^m(f_\psi) \quad \text{for} \quad r \leq \frac{1}{3}.$$

(ii) Let $f_0 \in \mathcal{C}(\psi)$. Then using Alexander's relation, $f_\psi(z) = zg'_\psi(z)$ it follows from first case that

$$B_N^m(f_0) \leq B_N^m(g_\psi) \quad \text{for} \quad r \leq \frac{1}{3}.$$

(iii) Let Ψ be convex and $f_0 \in \mathcal{F}(\Psi)$. Then from Lemma 2.2 , it follows that $f_0(z)/z \prec f_\Psi(z)/z$. This using [11, Lemma 2.1] gives that

$$B_N^m(f_0) \leq B_N^m(f_\Psi) \quad \text{for} \quad r \leq \frac{1}{3}.$$

Hence, it follows that

$$B_N^m(F, r) \leq L(r) \left(\frac{1 - r^\alpha}{1 - r}\right) < 1$$

for $|z| = r < \min\{1/3, r_0\}$, where r_0 is the minimum positive root of the polynomial equation

$$L(r)(1 - r^\alpha) + r - 1 = 0,$$

and $L(r)$ is as defined in the hypothesis.

□

Corollary 2.3.1. *Let $F(z)=\sum_{k=0}^{\alpha-1}\bar{z}^k f_k(z)$ be a poly-analytic function of order α, where f_k are analytic such that $g_k=f_0+\overline{f_k}$ is sense-preserving in Δ and $f_k(0)=0$ for each $k=0,\cdots,\alpha-1$. Let $\beta\in(0,3-2\sqrt{2}]$ and suppose that $f_0\in\mathcal{BS}(\beta)$. Then*

$$\sum_{n=1}^{m}\sum_{k=0}^{\alpha-1}|a_{n,k}|r^{k+n}<1,\quad \text{if}\quad |z|<1/3.$$

Corollary 2.3.2. *Let $F(z)=\sum_{k=0}^{\alpha-1}\bar{z}^k f_k(z)$ be a poly-analytic function of order α, where f_k are analytic such that $g_k=f_0+\overline{f_k}$ is sense-preserving in Δ and $f_k(0)=0$ for each $k=0,\cdots,\alpha-1$. Then*

$$B_N^m(F,r)=\sum_{n=N}^{m}\sum_{k=0}^{\alpha-1}|a_{n,k}|r^{k+n}<1,\quad \text{if}\quad |z|<\min\{1/3,r_0\}$$

where $N\in\mathbb{N}$ and r_0 is the minimal positive root of the equations:

$$\left(\frac{r}{1-r}\right)(r^N-r^m)(1-r^\alpha)+r-1=0,\quad \text{if}\quad f_0\in\mathcal{C}.$$

and

$$\left(\frac{r}{1-r}\right)\left[\frac{r^N-r^m}{r(1-r)}-(m-1)r^m+(N-1)r^N\right](1-r^\alpha)+r-1=0,\ \text{if}\quad f_0\in\mathcal{S}^*.$$

Remark 6. Let $F(z)=\sum_{k=0}^{\alpha-1}\bar{z}^k f_k(z)$ be a poly-analytic function of order α. Assume that $f_k\in\mathcal{C}$ for each $k=0,\cdots,\alpha-1$. Then the Rogosinski sum

$$B_1^m(F,r)=\sum_{n=1}^{m}\sum_{k=0}^{\alpha-1}|a_{n,k}|r^{k+n}<1\quad\text{in}\quad |z|<R_\alpha$$

where the Rogosinski radius R_α is the minimal positive root of the equations

$$r(1-r^\alpha)-(1-r)^2=0.$$

Here, we note that the radius $R_\alpha>1/3$.

Motivated from the Remark 6, we discuss the result when f_k comes from the Janowski-starlike class.

Theorem 2.4. *Let $F(z)=\sum_{k=0}^{\alpha-1}\bar{z}^k f_k(z)$ be a poly-analytic function of order α. Assume that $f_k\in\mathcal{S}^*[D,E]$ for each $k=0,\cdots,\alpha-1$. Then the Rogosinski sum*

$$B_1^m(F,r)=\sum_{n=1}^{m}\sum_{k=0}^{\alpha-1}|a_{n,k}|r^{k+n}<1\quad in\quad |z|<R$$

for each $m\geq 2$, where the Rogosinski radius R is the minimal positive root of the equations

$$f_j(r)(1-r^\alpha)+r-1=0,$$

where

$$f_j(z)=\begin{cases} z(1+Ez)^{\frac{D-E}{E}}, & E\neq 0;\\ ze^{Dz}, & E=0.\end{cases}\tag{2.12}$$

Proof. Since for each k, it is well-known that

$$|a_{n,k}|\leq\prod_{t=0}^{n-2}\frac{|E-D+Et|}{t+1},$$

where the equality is achieved for the function f_j given by (2.12). Now

$$\begin{aligned} B_N^m(F,r) &= \sum_{k=0}^{\alpha-1} r^k \left(\sum_{n=N}^{m} |a_{n,k}| r^n \right) \\ &\leq \sum_{k=0}^{\alpha-1} r^k \left(\sum_{n=N}^{m} \prod_{t=0}^{n-2} \frac{|E-D+Et|}{t+1} r^n \right) \\ &= \sum_{k=0}^{\alpha-1} r^k B_N^m(f_j,r) = B_N^m(f_j) \sum_{k=0}^{\alpha-1} r^k = B_N^m(f_j) \left(\frac{1-r^\alpha}{1-r} \right). \end{aligned}$$

Therefore, $B_N^m(F,r) < 1$ holds in $|z| < R_N^m$, where R_N^m is the minimum positive root of the polynomial equation

$$B_N^m(f_j,r)(1-r^\alpha) + r - 1 = 0.$$

Since $B_N^m(f_j) \leq B_N^{m+1}(f_j)$. Therefore, taking $N = 1$ and $m \to \infty$ shows that R_N^m approaches the Rogosisnki radius R. □

Corollary 2.4.1. *Let $F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)$ be a poly-analytic function of order α, where f_k are analytic such that $g_k = f_0 + \overline{f_k}$ is sense-preserving in Δ and $f_k(0) = 0$ for each $k = 0, \cdots, \alpha-1$. Then*

$$B_1^m(F,r) = \sum_{n=1}^{m} \sum_{k=0}^{\alpha-1} |a_{n,k}| r^{k+n} < 1, \quad \text{if} \quad |z| < \min\{1/3, r_0\}$$

where r_0 is the minimal positive root of the equations:

$$\left(\frac{r}{1-r} \right) (r - r^m)(1-r^\alpha) + r - 1 = 0, \quad \text{if} \quad f_0 \in \mathcal{C}.$$

and

$$\left(\frac{r}{1-r} \right) \left[\frac{r-r^m}{r(1-r)} - (m-1)r^m \right] (1-r^\alpha) + r - 1 = 0, \quad \text{if} \quad f_0 \in \mathcal{S}^*.$$

Remark 7. In Corollary 2.4.1, when $f_0 \in \mathcal{C}$. Then for each value of $m, \alpha \geq 2$,

$$\sum_{n=1}^{m} \sum_{k=0}^{\alpha-1} |a_{n,k}| r^{k+n} < 1 \quad \text{in} \quad |z| < \frac{1}{3}.$$

When $f_0 \in \mathcal{S}^*$, then $r_0 \leq 1/3$.

Corollary 2.4.2. *Let $F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)$ be a poly-analytic function of order α, where f_k are analytic such that $g_k = f_0 + \overline{f_k}$ is sense-preserving in Δ and $f_k(0) = 0$ for each $k = 0, \cdots, \alpha-1$. Then*

$$B_1^m(F,r) = \sum_{n=1}^{\infty} \sum_{k=0}^{\alpha-1} |a_{n,k}| r^{k+n} < 1, \quad \text{if} \quad |z| < \min\{1/3, r_0\}$$

where r_0 is the minimal positive root of the equations:

$$(1-r)^2 - r^2 + r^{\alpha+2} = 0, \quad \text{if} \quad f_0 \in \mathcal{C}$$

and

$$(1-r)^3 - r + r^{\alpha+1} = 0, \quad \text{if} \quad f_0 \in \mathcal{S}^*.$$

We now conclude this section with an interesting class studied by Bhowmik [7] on the basis of geometry, also see the family of Generalized Pascal Snails [15], which is as follows:

Definition 2. [7] A function $f : \Delta \to \mathbb{C}$ is said to be in the class of concave univalent functions with opening angle $\pi\beta$, $\beta \in [1, 2]$, at infinity if it fulfills the following conditions:

(a) f is analytic in Δ with the standard normalization and $f(1) = \infty$.
(b) f maps Δ conformally onto a set whose complement with respect to $\mathbb{C}$ is convex.
(c) the opening angle of $f(\Delta)$ at ∞ is less than or equal to $\pi\beta$.

For the function $f \in Co(\beta)$, the image domain $\mathbb{C}/f(\Delta)$ is closed and unbounded. Bhowmik [7] obtained sharp Coefficient's bounds, growth theorem, rotation theorem and distortion theorem etc. with the extremal function given by

$$f_\beta(z) := \frac{1}{2\beta}\left[\left(\frac{1+z}{1-z}\right)^\beta - 1\right]. \tag{2.13}$$

It seems interesting to study as an independent interest the class defined by

$$\mathcal{F}(f_\beta) := \left\{f \in \mathcal{A} : \left(\frac{zf'(z)}{f(z)} - 1\right) \prec f_\beta(z)\right\}.$$

For the class $Co(\beta)$, Bhowmik and Das [8] studied the Bohr phenomenon. Now we explore this in connection with poly-analytic functions.

Theorem 2.5. *Let $F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)$ be a poly-analytic function of order α, where $f_k \in Co(\beta)$. Then*

$$B_N^m(F, r) = \sum_{n=N}^{m}\sum_{k=0}^{\alpha-1} |a_{n,k}| r^{k+n} < 1, \quad \text{if} \quad |z| < r_\beta(N)$$

where $N \in \mathbb{N}$ and $r_\beta(N)$ is the minimal positive root of the equation

$$(1 - r^\alpha)B_N^m(f_\beta, r) + r - 1 = 0,$$

where f_β is given by (2.13).

Proof. First note that

$$\begin{aligned} B_N^m(F, r) &= \sum_{k=0}^{\alpha-1} r^k \left(\sum_{n=N}^{m} |a_{n,k}| r^n\right) \leq \sum_{k=0}^{\alpha-1} r^k \left(\sum_{n=N}^{m} A_n r^n\right) = \sum_{k=0}^{\alpha-1} r^k B_N^m(f_\beta, r) \\ &= B_N^m(f_\beta) \sum_{k=0}^{\alpha-1} r^k = B_N^m(f_\beta)\left(\frac{1-r^\alpha}{1-r}\right). \end{aligned}$$

Therefore,

$$B_N^m(F, r) < 1$$

holds in $|z| < r_\beta(N)$, where $r_\beta(N)$ is the minimum positive root of the polynomial equation

$$(1 - r^\alpha)B_N^m(f_\beta, r) + r - 1 = 0.$$

This completes the proof. □

Remark 8. Taking $\beta = 1$, in Theorem 2.5 yields the special case for convex functions.

From the Proof of Theorem 2.5, we deduce that

Corollary 2.5.1. *Let $F(z) = \sum_{k=0}^{\alpha-1} \bar{z}^k f_k(z)$ be a poly-analytic function of order α, where $f_k \in \mathcal{F}(f_1)$. Then*

$$|F(z)| + \sum_{n=N}^{\infty}\sum_{k=0}^{\alpha-1} |a_{n,k}| r^{k+n} < 1, \quad \text{if} \quad |z| < \min\{r_0, 1/3\}$$

where $N \in \mathbb{N}$ and r_0 is the minimal positive root of the equation

$$(1-r^{\alpha})(r+r^{N})-(1-r)^{2}=0.$$

As $\alpha \to \infty$ and $N=2$, then $r_0 \to 1/3$ and is best possible in this case.

3. Classical coefficient-functionals bounds for $\mathcal{F}(\psi)$

Suppose that the function ψ is univalent in Δ with $\psi(0)=0$. Then we can write

$$\psi(z)=A_1z+A_2z^2+A_3z^3+\cdots. \tag{3.1}$$

Further, if the function $1+\psi(z)$ belongs to the Carathéodory class then the Taylor's coefficients satisfies the well known bounds $|A_n|\leq 2$. Otherwise, when $A_1\neq 0$ then we have $|A_n|/|A_1|\leq n$ for each $n\in\mathbb{N}$. Since the Taylor's coefficients of the function $f(z)=z+\sum_{k=2}^{\infty}a_kz^k$ in the class $\mathcal{F}(\psi)$ can be expressed in the form of A_i, therefore to study the behavior of $|a_k|$ and some other coefficient functional, it is natural to consider the normalized form of f, that is $A_1=1$.

We start with some basic transformations which hold also for the class $\mathcal{S}$ of normalized univalent functions.

Remark 9. Let $f\in\mathcal{F}(\psi)$, where $\psi\in\mathcal{S}$. Then the following transformations belong to $\mathcal{F}(\psi)$:

(*i*) Let $\alpha\in\mathbb{R}$,

$$e^{i\alpha}f(e^{i\alpha}z)=z+\sum_{n=2}^{\infty}a_ne^{i(n-1)\alpha}z^n.$$

(*ii*) Let $0<t\leq 1$,

$$\frac{1}{t}f(tz)=z+\sum_{n=2}^{\infty}a_nt^{n-1}z^n.$$

(*iii*) Let k is an positive integer

$$[f(z^k)]^{\frac{1}{k}}=z+\frac{a_2}{k}z^{k+1}+\frac{1}{2k^2}(2ka_3-(k-1)a_2^2)z^{2k+1}+\cdots.$$

First, we obtain sharp initial coefficient's bounds, Zalcman [10, 27] and the Fekete-Szegö [3, 28] functional.

Theorem 3.1. *If $f\in\mathcal{F}(\psi)$, where $\psi\in\mathcal{S}$. Then*

(1) *sharp bounds for initial coefficient's are*

$$|a_2|\leq A_1,\quad 2|a_3|\leq\begin{cases}A_2+A_1^2, & \text{if}\quad A_1^2+A_2\geq A_1\\ A_1, & \text{if}\quad A_1\geq A_1^2+A_2\geq -A_1\\ -(A_2+A_1^2), & \text{if}\quad A_1^2+A_2\leq -A_1\end{cases}$$

and

$$|a_4|\leq\frac{A_1}{3}H(q_1,q_2),\quad \text{where}\quad\begin{cases}q_1=\frac{4A_2+3A_1^2}{2A_1}\\ q_2=\frac{2A_3+A_1^3+3A_1A_2}{2A_1}\end{cases}$$

(2) *sharp bound for the Zalcman functional of early coefficients is*

$$|a_2a_3-a_4|\leq\frac{A_1}{3}H(q_1,q_2),\quad \text{where}\quad\begin{cases}q_1=\frac{A_1^2+4A_2}{2A_1}\\ q_2=\frac{A_3-A_1^2}{A_1}\end{cases}$$

(3) *sharp bound for the Fekete-Szegö functional is*

$$|a_3-\nu a_2^2|\leq\begin{cases}\frac{1}{2}(A_2+(1-2\nu)A_1^2), & \text{if}\quad A_1+(2\nu-1)A_1^2\leq A_2\\ A_1/2, & \text{if}\quad \frac{A_2-A_1+A_1^2}{2A_1^2}\leq\nu\leq\frac{A_2+A_1+A_1^2}{2A_1^2}\\ -\frac{1}{2}(A_2+(1-2\nu)A_1^2), & \text{if}\quad (2\nu-1)A_1^2\geq A_1+A_2\end{cases}$$

where $H(q_1, q_2)$ *and* (q_1, q_2) *are given by* [32, Lemma 2].

Proof. The proof follows by employing the technique of [3, Theorem 1], a lemma of Ma and Minda [28] and [32, Lemma 2]. Here in this proof, we simply notice that $|A_3| \leq 3$, which is not the case for Ψ in the class $\mathcal{S}^*(\Psi)$. □

Corollary 3.1.1. *Let* $f \in \mathcal{BS}(\beta)$, $0 \leq \beta < 1$. *Then*

(1) *sharp bounds for initial coefficient's are*

$$|a_2| \leq 1, \quad |a_3| \leq 1/2$$

and

$$|a_4| \leq \frac{1}{3} H\left(\frac{3}{2}, \frac{2\beta+1}{2}\right)$$

(2) *sharp bound for the Zalcman functional of early coefficients is*

$$|a_2 a_3 - a_4| \leq \frac{1}{3} H(1/2, \beta - 1)$$

(3) *sharp bound for the Fekete-Szegö functional is*

$$2|a_3 - \nu a_2^2| \leq \begin{cases} (1-2\nu), & \text{if} \quad \nu \leq 0 \\ 1, & \text{if} \quad 0 \leq \nu \leq 1 \\ (2\nu - 1), & \text{if} \quad \nu \geq 1 \end{cases}$$

where $H(q_1, q_2)$ *is given by* [32, Lemma 2]. *The equality cases follow for the extremal function*

$$f_\beta(z) = \left(\frac{1 + z\sqrt{\beta}}{1 - z\sqrt{\beta}}\right)^{\frac{1}{2\sqrt{\beta}}} = z + z^2 + \frac{1}{2} z^3 + \frac{1}{3}\left(\beta + \frac{1}{2}\right) z^4 + \cdots .$$

Remark 10. In Corollary 3.1.1, part-1 improves $|a_3|$ and $|a_4|$ bounds in [16, Theorem 2.2], while part-3 reduces to [16, Corollary 2.2].

Corollary 3.1.2. *Let* $f \in \mathcal{S}_{cs}(\beta)$, $0 \leq \beta < 1$. *Then*

(1) *sharp bounds for initial coefficient's are*

$$|a_2| \leq 1, \quad 2|a_3| \leq 2 - \beta$$

and

$$|a_4| \leq \frac{1}{3} H\left(\frac{7 - 4\beta}{2}, \frac{1 + \beta^2}{1 + \beta} + \frac{4 - 3\beta}{2}\right)$$

(2) *sharp bound for the Zalcman functional of early coefficients is*

$$|a_2 a_3 - a_4| \leq \frac{1}{3} H\left(\frac{5 - 4\beta}{2}, \beta(\beta - 1)\right)$$

(3) *sharp bound for the Fekete-Szegö functional is*

$$2|a_3 - \nu a_2^2| \leq \begin{cases} 2(1-\nu) - \beta, & \text{if} \quad 1 - 2\nu \geq \beta \\ 1, & \text{if} \quad 1 - \beta \leq 2\nu \leq 3 - \beta \\ \beta - 2(1-\nu), & \text{if} \quad 2\nu \geq 3 - \beta \end{cases}$$

where $H(q_1, q_2)$ *is given by* [32, Lemma 2]. *The equality cases follow for the extremal function*

$$f_\beta(z) = z\left(\frac{1 + \beta z}{1 - z}\right)^{\frac{1}{1+\beta}} = z + z^2 + \left(1 - \frac{\beta}{2}\right) z^3 + \frac{1}{6}(6 - 5\beta + 2\beta^2) z^4 + \cdots .$$

Using the coefficients of the analytic function f given by $f(z) = z + \sum_{k=n+1}^{\infty} b_k z^k$, Pommerenke, and Noonan and Thomas considered the Hankel determinant $H_q(n)$ which is defined by

$$H_q(n) := \begin{vmatrix} b_n & b_{n+1} & \dots & b_{n+q-1} \\ b_{n+1} & b_{n+2} & \dots & b_{n+q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n+q-1} & b_{n+q} & \dots & b_{n+2(q-1)} \end{vmatrix},$$

where $b_1 = 1$. Estimation of upper bound of $|H_3(1)|$ [21, 23], $|H_2(2)|$ and $|H_2(1)|$ [3] for the functions belonging to various subclasses of univalent functions is a usual phenomenon in GFT, see [19]. For results on $\mathcal{S}^*(\Psi)$, we refer to see [3, 4, 24, 27]

We next tackle second hankel determinant $H_2(2) = (a_2a_4 - a_3^2)$ for the class $\mathcal{F}(\psi)$: Here one should note that we have $|A_3| \leq 3$, which is not the case in Ma-Minda function Ψ.

Theorem 3.2. *If $f \in \mathcal{F}(\psi)$, where $\psi \in \mathcal{S}$.*

(*i*) *If $|A_2| \leq A_1$ and $|A_1^4 - 4A_1A_3 + 3A_2^2 + 6A_1^2A_2| \leq 3A_1^2$ holds , then*

$$|H_2(2)| \leq A_1^2/4.$$

(*ii*) *If either*

$$|A_2| \geq A_1 \quad \textit{and} \quad |A_1^4 - 4A_1A_3 + 3A_2^2 + 6A_1^2A_2| \geq A_1|A_2| + 2A_1^2,$$

or the conditions

$$|A_2| \leq A_1 \quad \textit{and} \quad |A_1^4 - 4A_1A_3 + 3A_2^2 + 6A_1^2A_2| \geq 3A_1^2$$

holds, then

$$|H_2(2)| \leq \frac{|A_1^4 - 4A_1A_3 + 3A_2^2 + 6A_1^2A_2|}{12}$$

(*iii*) *If $|A_2| > A_1$ and $2A_1|A_2| + 2A_1^2 \geq |A_1^4 - 4A_1A_3 + 3A_2^2 + 6A_1^2A_2|$ holds, then*

$$|H_2(2)| \leq \frac{A_1^2}{4} - \frac{4A_1^2(|A_2| - A_1)^2}{48(|A_1^4 - 4A_1A_3 + 3A_2^2 + 6A_1^2A_2| - 2A_1|A_2| - A_1^2)}.$$

Proof. Let us consider the following series expansion:

$$\frac{zf'(z)}{f(z)} = 1 + a_2z + (2a_3 - a_2^2)z^2 + \cdots.$$

Since

$$\frac{zf'(z)}{f(z)} - 1 = \psi(\omega(z)) = \psi\left(\frac{p(z)+1}{p(z)-1}\right).$$

Therefore, we get

$$2a_2 = A_1p_1, \quad 8a_3 = A_2p_1^2 - A_1(p_1^2 - 2p_2) + A_1^2p_1^2$$

and

$$\begin{aligned} a_4 =& \frac{1}{48}\Big(2p_1(-2A_2p_1^2 + A_3p_1^2 + 4A_2p_2) + 3A_1^3p_1^3 - A_1^2p_1(p_1^2 - 2p_2) \\ &+ A_1\Big(p_1^3(6 + 2(A_2 - 1)) + 3(A_2 - 1) + 1 - 2A_2\Big) - 8p_1p_2 + 8p_3\Big) \end{aligned}$$

Now using the expressions from Lemma for p_2 and p_3 gives

$$\begin{aligned} a_2a_4 - a_3^2 =& \frac{1}{192}\Big(Mp^4 + 2p^2x(4 - p^2)A_1A_2 - 4p^2x^2(4 - p^2)A_1^2 \\ &- 3(4 - p^2)^2x^2A_1^2 + 8p(4 - p^2)z(1 - |x|^2)A_1^2\Big), \end{aligned} \tag{3.2}$$

where $M := -A_1^4 + 4A_1A_3 - 3A_2^2 - 6A_1^2A_2$. Assume $p_1 = p \in [0,2]$ because the function $P(e^{i\theta}z)$ is Carathéodory for each $0 \leq \theta < 2\pi$ if $P(z)$ is so. Now applying the triangle inequality in (3.2) with $|x| = \rho$ gives

$$\begin{aligned}|a_2a_4 - a_3^2| &\leq \frac{1}{192}\Big(|M|p^4 + 4p^2(4-p^2)\rho^2A_1^2 + 8A_1^2p(4-p^2)(1-\rho^2) \\ &\quad + 2A_1|A_2|p^2(4-p^2)\rho + 3A_1^2(4-p^2)^2\rho^2\Big) \\ &=: G(p,\rho).\end{aligned}$$

It is an easy exercise to note that $\max_{0\leq\rho\leq1} G(p,\rho) = G(p,1) =: F(p)$. Further, a simplication gives

$$\begin{aligned}192F(p) &= \left(|M| - A_1(3xA_1 - 2|A_2| - 4A_1)\right)p^4 + 8A_1\left(|A_2| - A_1\right)p^2 + 16A_1^2 \\ &=: \left(Ap^4 + Bp^2 + C\right).\end{aligned}$$

The result follows by applying the known inequality

$$\max_{0\leq t\leq 4}\left(At^4 + Bt^2 + C\right) = \begin{cases} C, & \text{if} \quad B \leq 0, 4A \leq -B \\ 16A + 4B + C, & \text{if} \quad B \geq 0, 8A \geq -B, \quad \text{or} \quad B \leq 0, 4A \geq -B \\ \frac{4AC-B^2}{4A}, & \text{if} \quad B > 0, 8A \leq -B. \end{cases}$$

This completes the proof. □

Remark 11. It is worth here to mention that the conclusion of the Theorem 3.2 also holds for the class $\mathcal{S}^*(\Psi)$ of Ma-Minda starlike function when $\Psi(z) := 1 + \psi(z)$ is a Carathéodory function and ψ has the form (3.1).

Corollary 3.2.1. *Let $f \in \mathcal{BS}(\beta)$, $0 \leq \beta < 1$. Then for all β*

$$|H_2(2)| = |a_2a_4 - a_3^2| \leq \frac{1}{4}.$$

The inequality is sharp with the extremal function

$$f(z) = z\exp\left(\int_0^z \frac{t}{1-\beta t^4}dt\right) = z + \frac{z^3}{2} + \frac{z^5}{8} + \left(\frac{1+8\beta}{48}\right)z^7 + \cdots.$$

Corollary 3.2.2. *Let $f \in \mathcal{S}_{cs}(\beta)$, $0 \leq \beta < 1$. Then*

$$|H_2(2)| = |a_2a_4 - a_3^2| \leq \frac{1}{4}, \quad \textit{if} \quad |6 - 8\beta - \beta^2| \leq 3$$

and

$$|H_2(2)| \leq \frac{|6 - 8\beta - \beta^2|}{12}, \quad \textit{if} \quad |6 - 8\beta - \beta^2| \geq 3.$$

The equality in first case holds with the extremal function

$$f(z) = z\exp\left(\int_0^z \frac{t}{1 + (\beta-1)t^2 - \beta t^4}dt\right) = z + \frac{z^3}{2} + \left(\frac{3-2\beta}{8}\right)z^5 + \cdots.$$

Further, application of this section can be seen for the non-Carathéodory regions given by Generalized Pascal Snails [15].

References

[1] ABDULHADI, Z.— EL HAJJ, L.: *On the Univalence of Poly-analytic Functions*, Comput. Methods Funct. Theory **22** (2022), 169–181.

[2] AIZENBERG, L.: *Remarks on the Bohr and Rogosinski phenomena for power series*, Anal. Math. Phys. **2** (2012), no. 1, 69–78.

[3] ALI, R. M.— RAVICHANDRAN, V.—SEENIVASAGAN, N.: *Coefficient bounds for p-valent functions*, Applied Mathematics and Computation **187** (2007), 35–46.

[4] ALARIF, N. M.—ALI, R. M.—RAVICHANDRAN, V.: *On the second Hankel determinant for the kth-root transform of analytic functions*, Filomat **31**(2) (2017), 227–245.

[5] ALKHALEEFAH, S. A.—KAYUMOV, I. R.—PONNUSAMY, S.: *Bohr-Rogosinski inequalities for bounded analytic functions*, Lobachevskii J. Math. **41** (2020), no. 11, 2110–2119.

[6] BOHR, H.: *A Theorem Concerning Power Series*, Proc. London Math. Soc. **13**(2) (1914), 1–5.

[7] BHOWMIK, B.: *On Concave univalent Functions*, Math. Nachr. **285** (2012), No. 5-6, 606–612.

[8] BHOWMIK, B.—DAS, N.: *Bohr phenomenon for subordinating families of certain univalent functions*, J. Math. Anal. Appl. **462** (2018), no. 2, 1087–1098.

[9] CHEN, H.—GAUTHIER, P.—HENGARTNER, W.: *Bloch constants for planar harmonic mappings*, Proc. Am. Math. Soc., **128** (2000), 3231–3240.

[10] CHO, N. E.—KWON, O. S.—LECKO, A.—SIM, Y. J.: *Sharp Estimates of Generalized Zalcman Functional of Early Coefficients for Ma-Minda Type Functions*, Filomat **32** 18 (2018), 6267–6280.

[11] GANGANIA, K.—KUMAR, S. S.: *Bohr-Rogosinski phenomenon for $\mathcal{S}^*(\psi)$ and $\mathcal{C}(\psi)$*, Mediterranean Journal of Mathematics, (2022). https://doi.org/10.1007/s00009-022-02074-7

[12] GANGANIA, K.—KUMAR, S. S.: *On Certain Generalizations of $\mathcal{S}^*(\psi)$*, Comput. Methods Funct. Theory (2021), https://doi.org/10.1007/s40315-021-00386-5. arXiv:2007.06069v1

[13] GANGANIA, K.—KUMAR, S. S.: *Bohr radius for some classes of Harmonic mappings*, Iran J Sci Technol Trans Sci (2022), https://doi.org/10.1007/s40995-022-01304-7

[14] GANGANIA, K.—KUMAR, S. S.: *A generalized Bohr-Rogosinski phenomenon*, (2022). https://arxiv.org/abs/2202.02560.

[15] KANAS, S.—MASIH, V. S.: *On the behaviour of analytic representation of the generalized Pascal snail*, Anal. Math. Phys. **11** 77 (2021). https://doi.org/10.1007/s13324-021-00506-3

[16] KARGAR, R.—EBADIAN, A.—SOKÓŁ, J.: *On Booth lemniscate and starlike functions*, Anal. Math. Phys. **9** (2019), 143–154.

[17] KAYUMOV, I. R.—KHAMMATOVA, D. M.—PONNUSAMY, S.: *Bohr-Rogosinski phenomenon for analytic functions and Cesáro operators*, J. Math. Anal. Appl. **496** (2021), no. 2, 124–824, 17 pp.

[18] KUMAR, S.—SAHOO, S. K.: *A Generalization of the Bohr-Rogosinski Sum*, Priprint, arXiv:2106.06502v1

[19] KUMAR, S. S.—KAMALJEET, G.: *A cardioid domain and starlike functions*, Anal. Math. Phys. **11** (2021), no. 2, 54.

[20] KUMAR, S. S.—KAMALJEET, G.: *On Geometrical Properties of Certain Analytic functions*, Iran. J. Sci. Technol. Trans. Sci. **45** (2021), 1437–1445.

[21] KWON, O. S.—LECKO, A.—SIM, Y. J.: *The bound of the Hankel determinant of the third kind for starlike functions*, Bull. Malays. Math. Sci. Soc. **42**(2) (2019), 767–780.

[22] LANDAU, E.—GAIER, D.: *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, third edition, Springer-Verlag, Berlin, (1986).

[23] LECKO, A.—SIM, Y. J.—ŚMIAROWSKA, B.: *The sharp bound of the Hankel Determinant Of the Third Kind for Starlike Functions of Order half*, Complex Anal. Oper. Theory. (2018).

[24] LEE, S. K.—RAVICHANDRAN,V.—SUPRAMANIAM, S.: *Bounds for the second Hankel determinant of certain univalent functions*, J. Ineq. Appl. 2013, 281 (2013), 1–17.

[25] LI, M.—SUGAWA, T.: *A note on successive coefficients of convex functions*, Comput. Methods Funct. Theory **17**(2) (2017), 179–193.

[26] LIU, M. S.: *Estimates on Bloch constants for planar harmonic mappings*, Sci. China Ser. A Math **52** (2009), 1142–1146.

[27] MA, W. C.: *Generalized Zalcman conjecture for starlike and typically real functions*, J. Math. Anal. Appl. **234**(1) (1999), 328-339.

[28] MA, W. C.—MINDA, D.: *A unified treatment of some special classes of univalent functions*, Proceedings of the Conference on Complex Analysis, Tianjin, Conf Proc Lecture Notes Anal, I Int Press, Cambridge, MA, 157–169 (1992).

[29] MASIH, V. S.—EBADIAN, A.—YALÇIN, S.: *Some properties associated to a certain class of starlike functions*, Math. Slovaca **69** (2019), 1329–1340.

[30] PIEJKO, K.—SOKÓŁ, J.: *Hadamard product of analytic functions and some special regions and curves*, J. Inequal. Appl. 2013 Art. 420, 13 pp.
[31] PIEJKO, K.—SOKÓŁ, J.: *On Booth lemniscate and hadamard product of analytic functions*, Math. Slovaca **65** (2015),1337–1344.
[32] PROKHOROV, D. V.—SZYNAL, J.: *Inverse coefficients for* (α, β)*-convex functions*, Annales Universitatis Mariae Curie-Sklodowska, Lublin-Polonia, Sectio A **35**(15) (1981), 125–143.
[33] ROGOSINSKI, W.: *Über Bildschranken bei Potenzreihen und ihren Abschnitten*, Math. Z. **17** (1923), no. 1, 260–276.
[34] RUSCHEWEYH,S.—STANKIEWICZ, J.: *Subordination under convex univalent functions*, Bull. Polish. Acad. Sci. Math. **33** (1985),499–502.
[35] SCHUR, I.—SZEGÖ, G.: *Über die Abschnitte einer im Einheitskreise beschränkten Potenzreihe*, Sitz.ber. Preuss. Akad. Wiss. Berl. Phys.-Math. Kl. (1925), 545–560.

Department of Applied Mathematics
Delhi Technological University
New Delhi
India
Email address: gangania.m1991@gmail.com

SAHARA PUBLIC LIBRARY
Publisher of Scientific Journals

NeuroPharmac Journal
Journal Homepage: https://www.neuropharmac.com

NeuroPharmac

Review

Therapeutic applications of mesenchymal cells derived from human placenta A comprehensive meta-review

Ritika Luthra[1], Saanya Yadav[1], Smita Rastogi Verma[1]*
[1]Department of Biotechnology, Delhi Technological University, Delhi

ARTICLE INFO

Article history
Received 11 November 2021,
Revised 29 November 2021,
Accepted 8 December 2021,
Published online
30 December 2021

Keywords
Placental mesenchymal stem cell (PMSC), cell therapy, clinical trials, regenerative medicine, multipotent.

ABSTRACT

There has been an increase in the incidence of diseases and injuries which has encouraged the advancements in treatments that could repair tissues as well as minimize the dependence on organ transplantation and tissue loss. Regenerative medicine based on stem cells is a newer integrative domain that has the potential to facilitate the regeneration of diseased and damaged tissues and organs. Stem cells can be effectively exploited for medical purposes with no severe challenges, according to a wide range of literature as well as long-term human and animal studies conducted on them. Amongst several types of stem cells, placenta-derived mesenchymal stem cells possess unique immunomodulatory characteristics as well as the potential to differentiate into various cell types, which make them suitable candidates for cellular therapies for many chronic diseases including cancers, heart and liver disorders, ulcers, bone damage, and neurological diseases, etc. The present study thus aims at providing a comprehensive overview of placenta-derived mesenchymal stem cells, their types, and potentially therapeutic for various chronic and non-chronic diseases. Recent pre-clinical and clinical studies conducted on them have also been incorporated.

Introduction

In recent times, modern science and medicine have seen an appreciable rise in stem cell research which has garnered recognition globally attributed to its enormous therapeutic potential in the field of regenerative medicine.[1] This is a result of the regenerative capability as well as the pliability possessed by the stem cells. These stem cells manifesting therapeutic potential originate and develop from embryonic as well as adult tissues. Nonetheless, ethical concerns along with the challenges in procuring them and tumorigenicity limit the clinical usage of embryonic stem cells. The adult stem cells have been discovered in various tissues and organs, such as bone marrow[2], brain, placenta[3], skin, adipose tissue le, umbilical cord[4], etc. However, the amount of stem cells obtained from adult tissues is limited. There are numerous risks involved in isolating them and these cells possess a minimal range of differentiation and proliferation after their extraction from the body. As a result, it is burdensome to obtain large quantities

*Corresponding author at: Department of Biotechnology, Delhi Technical University, Delhi-110042, India
E-mail address: smitar@dtu.ac.in
https://orcid.org/0000-0002-3812-5314
https://doi.org/10.37881/1.637

of stem cells from them. On the other hand, a good source of mesenchymal stem cells (MSCs) in the human placenta, a vital part of the body found at the time of pregnancy for assisting fetal development. It is a multicellular immunoregulatory organ where both maternal, as well as fetal cells, exist side-by-side. It performs several functions including exchange of metabolites and regulation of the secretion of hormones between the mother and fetus along with the maintenance of immune tolerance between the two.[5] Placenta is also an abundant reservoir of stem cells in comparison to other sources like bone marrow, adipose tissue, etc. where there is a decline in the recuperation of cells with benefactor's age.[6] MSCs procured from the placenta reveal various features making them a better alternative for cellular therapy in comparison to their counterparts from adult tissues. The placenta is contemplated to be a medical waste having no ethical issues in its utilization as seen in the case of embryonic stem cells.[3] It is voluntarily removed at the time of birth without the requirement of any encroaching procedures unlike those needed in other adult stem cell sources. Because of their primordial origin, placental cells possess great ambidexterity as well as the ability to differentiate.[7] Stem cells procured from the placenta present a distinctive range of mesenchymal markers,[8,9] along with the potential to differentiate into neurogenic and glial cells[10], insulin-producing cells[11] as well as liver cells.[12] MSCs acquired from the placenta have been employed for treating a wide range of diseases, including cancer, cardiovascular, neurological, and liver diseases, ulcers, as well as bone and cartilage diseases. These MSCs are of comparatively recent types, exhibiting particular immunomodulatory functions with undetermined mechanisms. Certain soluble factors secreted by placental MSCs appear to be the primary cause for therapeutic effects. This shows that placenta-derived MSCs possess paracrine effects.[13] Contrastingly, because of their migration capabilities and tropism for wounded regions, placental MSCs can also act as mobile carriers for drugs.

Recently, there has been immense progress in the field of nanotechnology for curing damaged tissues or organs. Owing to the unique features manifested by placental MSCs, these cells together with nanotechnology will turn out to be a considerable and reassuring domain that may bestow remarkable contributions in the field of stem cell therapy in subsequent times. Therefore, this review seeks to incorporate the formation of placenta-derived mesenchymal stem cells and gives a comprehensive overview of their potential therapeutic applications. Furthermore, the prospects have also been discussed.

Sources of mesenchymal stem cells

MSCs may be extracted from a variety of tissues, including bone marrow, adipose tissue, and synovium, as well as human umbilical cord blood, and bone marrow is one of the most important sources of MSCs.[14] The human body contains adipose tissue (AT) in numerous places, including visceral and subcutaneous fat pads, as well as the pelvic region. AT-derived MSCs, also known as Adipose Stem Cells (ASCs), have comparable properties to Bone marrow-MSCs but, unlike Bone marrow mesenchymal stem cells (BM-MSCs), are easier to isolate in large quantities and have a larger expansile property. The physician must consider the difficulties of obtaining the samples as well as the potential negative consequences of harvesting the cells on the donor while choosing an appropriate cell source. Collecting MSCs from BM-MSCs, for example, might cause pain, hemorrhage, or infection, making it more difficult than harvesting cells from peripheral blood or surgical leftovers like AT or birth-derived tissues.[15] MSCs from peripheral blood cells can be obtained by density gradient centrifugation and a large volume of MSCs can be extracted as compared to bone marrow.[16] However, placental mesenchymal stem cells are harvested with minimal invasiveness thereby being the best candidate for therapeutics.[17]

Different forms of placental mesenchymal stem cells based on origin

Mesenchymal stem cells are derived from mesoderm.[18] Specifically, placental mesenchymal stem cells can be derived from amniotic membrane, amniotic fluid, chorionic plate as well as chorionic villus.[19]

The amniotic membrane is a placental component that develops in extra-embryonic tissue and provides additional nutrition to the fetus throughout pregnancy.[20] Cells derived from amniotic membranes are called amniotic membrane-derived MSCs (AMe-MSC). They are pluripotent. AMe-MSCs have been shown to reduce inflammation and improve immunomodulation. They induce remyelination, increase angiogenesis, reduce oxidative stress, and control matrix metalloproteinases.[21] Amniotic fluid-derived mesenchymal stem cells (AF-MSCs) are made up of differentiated and undifferentiated progenitor cells generated from fetal tissue. AF-MSCs are more pluripotent and less differentiated than BM-MSCs, having a longer replicative lifetime and proliferative capacity. In AF-MSCs, cytokines and chemokines stimulate vasculogenesis, angiogenesis, and osteogenesis due to their paracrine nature. *In vivo*, AF-MSCs do not produce teratoma.[22] Because of the ease of their acquisition, multipotency, low immune response, reduced donor damage, and acceptable ethical issue, AM-MSCs and AF-MSCs have been reported as a better new prospective field of regenerative medicine when compared to other MSC sources like chorionic, Wharton's jelly, etc.[23]

The amnion, extra-amniotic mesenchymal stem cells, cytotrophoblast, and syncytiotrophoblast make up the chorionic plate. Chorionic plate-derived MSCs (CP-MSCs) are extracted from chorionic placental plates. Compared to other placental stem cells, CP-MSCs have a greater ability to suppress T-cell proliferation and have a stronger angiogenic potential. Higher genes for differentiation into adipogenic, osteogenic, and hepatogenic lineages are expressed in them.[24] Chorionic villus cells, which are made up of stromal fibroblasts, endothelial cells, and macrophages, have characteristics similar to MSCs.[25] Chorionic villous-derived MSCs (CV-MSCs) are more pluripotent and have a slower aging phenotype than bone marrow-derived stem cells.[26]

Placenta and Stem Cell Therapy

Stem cell therapy, also known as regenerative medicine, emphasizes the repair and rejuvenation of injured tissues and organs which may occur due to age, disease, or any physical damage. Such approaches turn out to be beneficial in cases where biological processes are incompetent for responding during the emergence of acute as well as chronic disorders. They may necessitate the transplantation of stem cells which will restore the injured tissues, instigate body functions for repairing of tissues, or work as mobile carriers for delivering remedial agents such as genes, medications, or cytokines. Stem cells are considered exceptional in regenerative medicine because of their ability to restore as well as to differentiate to various specialized cells. Advancements in the field of regenerative medicine have drawn attention to the complex characteristics and heterogeneity of stem cells procured from distinct sources. In the past few years, researchers have traversed various properties of adult MSCs suggesting them to be advantageous in regenerative medicine.[27] Hence, MSCs obtained from several sources have displayed enormous prospects in damaged tissues as well as organ repairing and restoration. Being an immune-privileged organ, placental cells exhibit low immunogenicity *in vitro*[28]as well as *in vivo*[29] when xenotransplanted in animals displaying a normal immune response. Researchers have illustrated the viability of placenta-derived cells for allogeneic transplantations.[30]

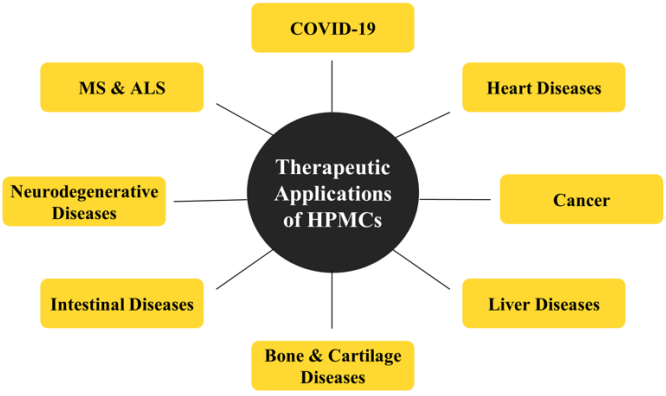
In stem cell therapy, along with tissue repairing, the short-lived paracrine actions are also affected by stem cells. The factors generated and released by stem cells induce these paracrine actions. These factors are responsible for controlling injuries, modulating the immune responses as well as promoting self-restoration in damaged tissues which are still alive.[31] Concerning modulation of the immune system, stem

cells obtained from the placenta possess an auxiliary advantage over those derived from other sources. This can be well explained by the primary role of the placenta in fetomaternal immune tolerance.[13] During pregnancy, fetal and, particularly, the tissues of the placenta are responsible for maternal tolerance. They provide immune-privileged conditions and regulate the immune system. MHC class II antigens which generally arbitrate transplant rejection are absent in placental cells.[32] Placental cells express fewer amounts of extremely diverse types of antigens belonging to MHC class I as well as the unconventional type HLA-G (histocompatibility antigen class I, G) which can be responsible for suppressing body defenses along with contributing to immune tolerance during pregnancy.[33,34] Additionally, placenta-derived cells diverge the maternal immune response towards immune tolerance via the secretion of soluble forms of MHC antigens, hormones, and cytokines.[35,36] Furthermore, B-cells as well as most of the T-cells vanish leaving behind regulatory T-cells, also called Tregs, as the larger population of T-cells that possess immunosuppressive as well as anti-inflammatory effects.[37]

Therapeutic applications of placenta-derived mesenchymal stem cells

Regenerative medicine is anticipated to bestow considerable advantages to individuals affected by a broad array of pathologies. Placenta-derived mesenchymal stem cells exhibit flexibility and pleiotropic characteristics which incorporate immune system modulation and control of inflammation. Some of the properties like angiogenesis, neuroprotective and anti-apoptotic are the properties that have been broadly assessed at preclinical levels.[7,38,39] Owing to these properties, HPMCs possess huge potential as therapeutics against several diseases, which are illustrated in Figure 1 and have been discussed in detail in further sections.

Figure 1: Therapeutic applications of human placenta-derived mesenchymal stem cells against various diseases



COVID-19

SARS-CoV-2 is classified as a novel coronavirus that belongs to the β–coronavirus family. After Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS), COVID-19 has emerged as the third zoonotic illness-based pandemic produced by viruses belonging to this family.[40] The genome of the COVID-19 virus is 26-32 kilobase pairs in length with crown-like projections on surfaces called spike proteins. The spike protein mainly interacts with the ACE-2 (angiotensin-converting enzyme) on alveolar epithelial cells and performs interspecies transmission of the genome.[41] This attachment leads to an elevated expression of ACE-2 which injures alveolar cells causing lesions. These lesions can cause serious systemic responses and eventually death if not treated on time. From affecting the lungs, these lesions can cause dysfunction of numerous organs.[42,43] An immune destruction in the form of cytokine

storm occurs in the body, in which pro-inflammatory cytokines, for example, interferons (INFs), interleukins (IL-1, 6, 12, 8, 33), tumor necrosis factors (TNFs), tumor growth factors (TGFs), accumulate in abundance.[44]

MSC treatment can inhibit the immune system from releasing a flood of cytokines and boost endogenous healing about the stem cell's reparative abilities. P-MSCs inhibit further infection of pulmonary alveolar cells by acting on the ACE-2 receptor and TMPRSS-2 (endothelial cell surface protein) receptor.[22] Placental MSC cells, when injected into the body, get entrapped in the lungs to restrict the systemic infusion of the virus. Thereafter, these MSCs have the potential to restore the required pulmonary environment, preserve alveolar epithelial cells, prevent pulmonary fibrosis, and treat lung dysfunction and COVID-19 pneumonia.[45] P-MSCs reduce the severity of cytokine storm caused by the virus, change the microenvironment to an anti-inflammatory state, and then further limit mononuclear cell entrance into the infected/damaged alveolar epithelium cells. In injured alveolar cells, P-MSCs increase immunomodulatory and anti-inflammatory properties. Thereby accelerating the repair mechanism of damaged alveolar cells.[22] Patients treated with MSC showed increased levels of peripheral lymphocytes and a decline in the amount of cytokine-secreting immune cells including CXCR3+ CD4+ T cells, CXCR3+ CD8+ T cells, and NK CXCR3+ cells compared with patients treated with conventional therapy.[45]

Myocardial Infarction

Cardiovascular disorders are a major cause of morbidity and death globally, with myocardial infarction (MI) being the most frequent cardiovascular disorder.[46] MI happens as a result of intervention in the circulation of blood to heart muscles which are accompanied by myocardial ischemia. This is influenced by the capability as well as the timing of the restoration of blood flow, and therefore, may or may not be reversible.[47] The impaired cardiac muscle is substituted by scar tissue as the restoration of the myocardium is practically not present, thereby limiting the functioning of the heart. Transplantation of placenta-derived mesenchymal stem cells acts as a promising approach for restoring heart function as well as lessening cardiac fibrosis because of the angiogenic and immunomodulatory characteristics possessed by them. These cells can potentially transform into myocardiocytes along with exhibiting voluntary movements *in vitro* environments indicating their therapeutic application in the process of heart repair.[7,48,49] Various researchers have inspected the influence of placenta-derived mesenchymal stem cells after transplanting them in animal models of MI. These stem cells when administered in the heart of rats after inducing MI exhibited incorporation into heart tissues and *in vivo* lineage programming to form myocardiocytes.[48] The prime route for arbitrating the relocation of mesenchymal stem cells in the vicinity of damaged tissues is the axis of chemokine receptor type 4 (CXCR4), and its ligand i.e., stromal cell-derived factor (SDF-1) forming CXCR4-SDF1. As observed, due to hypoxia, CXCR4 is instigated at a high extent in placental mesenchymal stem cells. Therefore, a huge chemotactic reaction is anticipated by the placental mesenchymal stem cells to the ischemic microenvironment of the damaged portion of the heart.[50] Injecting these stem cells intravenously in an infarcted rat manifested an upraised heart function after around 32 weeks from infarction.[51] Provisioning these stem cells with a novel ester of hyaluronan linked to butyric as well as retinoic acid (HBR) reinforces their restorative capability. Transplanting these provisioned stem cells in pigs generated a notable lowering in the extent of scar, more myocardial transmigration and glucose absorption, elevated density of capillaries, as well as a decline of fibrous tissues.[52] Researchers have also assessed the angiogenic paracrine effect of the conditioned medium of placental mesenchymal stem cells. In a rat model of reperfusion injury, injecting the conditioned medium of these stem cells reduced the size of infarction as well as myocardiocyte apoptosis, whilst encouraging the capillary density in the borderline of infarction.[53]

Intracoronary injection of umbilical cord mesenchymal stem cells in people suffering from MI led to a healthy and markedly increased myocardial tolerance and circulation in the infarcted area. Enhancements in certain factors, like increased left ventricular ejection fraction as well as decreased end-diastolic volumes and LV end-systolic volumes, have been reported up until 18 months after therapy.[54] A phase I/II clinical study named RIMECARD showed the security and effectiveness of these stem cells which were injected intravenously in people suffering from chronic heart failure. There was also a decreased fraction ejection. Advancements in left ventricular activity, cognitive function, and life expectancy have been identified in treated groups.[55] The capacities of first trimester placental chorionic mesenchymal stem cells (FCMSCs) and third trimester placental chorionic mesenchymal stem cells (TCMSCs) have also been studied. FCMSCs had a better capacity to produce EC differentiation, as shown by improved *in vitro* morphology, angiogenic ability, *in vivo* cardiac function, and higher levels of pro-angiogenesis gene expression than TCMSCs.[56]

Critical Limb Ischemia

Critical limb ischemia is the most dangerous and advanced phase of peripheral artery disease which results from the thickening of walls of arteries due to the formation of plaque and brings scarcity in the blood perfusion, incorporating acute limb ischemia, gangrene, and even ulcers.[57,58] It is accompanied by accelerating stenosis, and eventually leads to blockage of peripheral arteries. The risks of developing peripheral artery disease increase with advancing age, high lipid content such as cholesterol (hyperlipidemia), high blood pressure, and predominantly diabetes. The hostile tissue environment generated by diabetes and ischemia can impair the functioning of a patient's cell products for autologous and allogeneic related therapies.[59] Angiogenic cells can contribute directly to the creation of new arteries while also providing endogenous growth factors, encouraging vascular expansion in a paracrine manner. Angiogenic treatment includes the use of angiogenic growth factors (HIF-1α, VEGF, FGF1, etc.), gene transfer techniques utilizing viral or non-viral vectors to transport a gene coding for a therapeutic protein to target tissues, or the use of angiogenic stem cells.[60] However, in several instances, amputation is the sole alternative possible for treating critical limb ischemia. This is because the blood capillaries cannot be rectified, and recompression of vessels is generated. It has been outlined in the preclinical trials that cell therapy exhibits several benefits in neovascularization in numerous mice suffering from hindlimb ischemia. Placental mesenchymal stem cells have illustrated pro-angiogenic properties, when introduced into the ischemic area of the infected limb via intramuscular injections, thereby enhancing the blood circulation as well as encouraging the development of new vessels.[24,61,62] PMSCs can be administered with insulin (from external sources) and even without insulin. Upon administration, the efficacy of PMSC was calculated. The results were promising as there were newly formed capillaries, a higher number of arterioles, and even high secretion of different proangiogenic factors.[63] Furthermore, conditioned medium from placental mesenchymal stem cells also possessed pro-angiogenic effects in a mouse model having hindlimb ischemia, equivalent to the transplanted placental stem cells in the same research, disclosing that action of placental mesenchymal stem cell led fundamentally from the paracrine effect of the angiogenic elements secreted from these stem cells.[24] Nonetheless, in one more examination, cells were more effective in comparison to the cell lysate in liberating blood circulation, presumably stipulating the significance of elongated paracrine action for optimal blood flow restoration.[63]

Stroke

Ischemic stroke, also called a cerebrovascular accident, is acute damage in the central nervous system caused due to compromised vasculature or brain perfusion, incorporating cerebral infarction, cerebral

bleed, and subarachnoid hemorrhage. It predominantly causes disability globally, and about 85% of strokes appear to be ischemic.[64] The frequently utilized curative technique for this condition is thrombolysis; however, many patients avoid out of window visits which is necessary for efficacious treatment. It has been outlined via experimental facts that regenerative medicine can lessen the degeneration of neurons as well as boost functional outcomes. Placental mesenchymal stem cells exhibit neuroprotective effects, which have been illustrated in stroke rat models. Administering these stem cells intravenously leads to a noteworthy enhancement of functional outcomes along with a considerable reduction in the mass of lesions, after 4 hours in the injury. This can be correlated with an escalation in the amounts of vascular endothelial and hepatocyte growth factors, as well as a brain-derived neurotrophic factor in the ischemic brain in comparison with the controls.[65] Similarly, stem cells can limit the injury caused by stroke. For the first time, Barzegar et al, (2021)[66] discovered that intravenous injection of placental MSCs administered at the moment of reperfusion dramatically reduced ischemia damage in the ipsilateral hemisphere. This protection is linked to a substantial return of normal blood flow to the brain following MCAO (middle cerebral artery occlusion).

Cancer

Over the years, our understanding and ability to deal with cancer have improved significantly. Despite the accelerated progress in the oncology domain, cancer remains a major lethality affecting both developing and developed parts of the world. The conventional treatment of cancer including surgical removal of the tumor, radiotherapy, and chemotherapy, have proven to be effective, however, severe side effects and drug resistance highlight the need to find an alternate therapy that could minimize the toxicity and still be efficient in treating the tumor cells.[67]

Placenta-derived mesenchymal stem cells offer a novel approach to cancer treatment. Placenta-derived MSCs have an intrinsic nature to migrate to the sites in the body which shows any kind of wear and tear. This property of PMSCs is majorly exploited in cancer treatment. Cancer and tumor sites behave as an unhealed wounded environment which sends a signal through paracrine signaling to PMSCs to migrate to the site of action. Inflammation is another factor that contributes to the movement of the cells.[68,69] This migration property was first observed in a xenograft mouse model.[70] PMSCs can be used as cellular Trojan horses as well for the targeted drug delivery. Nanoparticles are a novel approach for drug delivery but being foreign particles, nanoparticles are prone to high immunogenic response and have a high rate of failure as well due to inefficient dissemination in tumors. All these can be overcome by using stem cells as their delivery system.[71,72] This compound system of drug delivery has proven to be quite efficient in tumor apoptosis. MSC cell membrane coated with doxorubicin enclosed in porous silica nanoparticles is a fine example.[73] Furthermore, PMSCs can also be programmed genetically to produce certain antitumor compounds and peptides. Generally, PMSCs are modified with viral particles for the expression of certain cytokines like interferon-β (INF-β). A study conducted by Shen et al. showed induced apoptosis in breast cancer cells when umbilical cord MSCs were subjected to adenovirus particles expressing INF-β.[74] They can also be modified for producing particular enzymes as evident in the case of cytosine deaminase-producing MSCs when injected into the brain hindered the growth of glioblastoma.[70] Endostatin, a neurovascular and anti-angiogenesis inhibitor, has been successfully administered using hPMSCs, which works as a gene vector in the treatment of ovarian cancer.[75] Along with acting as an accessory to other forms of therapies, PMSCs as their own also have anti-tumorigenic responses. Studies have demonstrated MSCs secreting cytotoxic peptides, mediating pro-apoptotic pathways, and inducing an anti-cancer immune response. Growth of breast cancer was attenuated by the placental MSCs through inhibiting ERK1/2 and activating the AKT pathway, leading to apoptosis.[76] TNF-related apoptosis-inducing ligand

(TRAIL) is another cytotoxic agent released by MSCs inducing apoptotic pathways, all leading to inhibiting the growth of several types of carcinomas.[77] Combining PMSCs with drugs to improve restrictions in treatment is also considered. Therapeutic medicines combining anti-angiogenic treatment and chemotherapy can be utilized with decidua-derived MSCs. The decidua mesenchymal stem cells were used to carry out a combination of endostatin transfection and doxorubicin-nanoparticle loading.[78]

Parkinson's Disease

In Parkinson's disease (PD) a gradual impairment of motor skills can be observed. This happens when neurons in the brain controlling the locomotive functions become functionally inactive or die. It is characterized by the loss of dopamine-releasing neurons and thereby, decreased levels of dopamine in the striatum. There is no cure available for the full-fledged treatment of the disease, however, few medicines and therapies have been employed to target the symptoms. Anticholinergics, dopamine agonists, and neuroprotective (drug therapy) are the most common treatments for PD to date.[14,79] When talking to stem cells, the main objective becomes to re-establish the dopaminergic neurons and related transmission signals. Neural progenitors produced by the differentiation of PMSCs were transplanted into a rat with PD. The results showed alleviation of the PD symptoms and the neural progenitors transformed into dopaminergic differentiation.[80] In a study on mice, MSCs can lower the amount of α-synuclein, block microglia cell polarization, and increase neuron survival.[81] It is believed and being studied that treatment with PMSCs can also help in symptoms unrelated to the motor like constipation and hyposmia through the secretion of anti-apoptotic and anti-inflammatory factors.

Alzheimer's Disease

Alzheimer's disease (AD) is yet another progressive disease that is characterized by dementia i.e., memory loss with aging. It is a result of abnormal accumulation of amyloid peptides and tau tangles all over the brain causing loss of neural activity.[79] Structurally, it usually starts with the hippocampus and entorhinal cortex, parts primarily responsible for memory response. With time, shrinkage in other parts of the brain is also visible. AD mice models injected with PMSCs demonstrated higher levels of amyloid breaking enzymes and also lower levels of pro-inflammatory peptides. PMSCs showed a strong immunomodulatory effect by increasing the levels of anti-inflammatory cytokines like IL-10 and TGF-β. Moreover, enhanced memory response was also observed.[82] MSCs generated from the placental amniotic membrane can help AD model mice enhance their spatial learning and memory. The number of amyloid plaques was also decreased, suggesting that MSC therapy generated from the human placental amniotic membrane might ameliorate AD pathogenesis and cognitive performance by regulating oxidative stress.[83]

Amyotrophic Lateral Sclerosis

Another neurological disease is amyotrophic lateral sclerosis (ALS), which is characterized by symptoms like muscle weakness, paralysis, respiratory problems ultimately leading to death. These symptoms are caused by the degeneration of neurons in the brain and spinal cord. Although no conclusive therapies are available for ALS, the stem cells put forward an approach that leads to extended lifespan. ALS mice models injected with PMSCs demonstrated delayed degeneration of motor neurons by shielding them from inflammatory peptides.[84] MSCs have been found to migrate to the spinal cord of mice and inhibit microglial activation and tissue glial proliferation, increasing the number of motor neurons, suggesting that MSCs may have neuroprotective properties. Experiments have demonstrated that exosomes produced from ASC can protect motor neuron-like NSC-34 cells from oxidative damage and improve cell viability.[14]

Multiple Sclerosis

The pathogenicity of multiple sclerosis (MS) can be divided into two sections; inflammatory immune response and the degenerative process. The main characteristic of MS is the demyelinated portion in the central nervous system. The demyelination is healed by the formation of a glial scar. MS is also an auto-immune disorder mediated by the T-cells, attacking the myelin sheath peptides leading to loss of neurons and dendrocytes. Therapies available and in the study mostly target the immune system. PMSCs provide a unique approach to the treatment of MS patients by regulating both the immune response and facilitating the regeneration of the neural cells. Many studies have been conducted which use PMSCs on autoimmune encephalomyelitis (EAE) mouse models. The symptoms of EAE closely resemble MS symptoms. An immune response is triggered by the injection of myelin antigen and transplantation of PMSCs at 5th day and 14th day from immunization was concluded to reduce disease lethality and extended survival.[85] Injection of PMSCs is the EAE mice models were also an effective approach as these delayed the motor symptoms, modulated immune response by inhibiting the T cell proliferation and reducing inflammatory reactions, preventing neural network loss, and decreasing the seriousness of MS.[86,87]

The therapeutic efficacy of transplanting these stem cells in patients suffering from MS has been tested in various trials. Injecting umbilical cord-derived mesenchymal stem cells intravenously comes out to be feasible and safe in such patients. Additionally, the overall effects on treatment groups were observed to be stable or increased in comparison to the standard group.[88] In another research, patients having relapsing-remitting MS or secondary progressive MS were spontaneously administered with placental stem cells (PDA-001, a preparation of hPMSCs) and several individuals who were treated had steady or declining expanded disability status scale ratings (EDSS), which showed potential reparative influence on the patients.[89]

Liver diseases

Cirrhosis is the final stage of many liver injuries like viral infections, metabolic disorders, alcoholism, and even acute liver failure resulting from various processes. The typical functioning of the liver is disturbed due to the formation of the scar by the extracellular matrix. It further leads to necroinflammation and fibrogenesis.[90,91] Liver cirrhosis is an irrevocable condition that can be deadly, and the sole option for curing the disease is liver transplantation. However, the lesser availability of donors along with the constant requirement of immunosuppressants restrict liver transplantation and this is why cell transplantation emerges to be an appropriate alternative. Besides fetal and adult hepatocytes, stem cells are also taken into consideration for cell transplantations. It has been outlined that PMSCs can potentially be beneficent because of their ability to transform to hepatic-like cells, thereby forming functioning 3D structures.[92] AF-MSCs can engraft in the injured liver and produce Annexin-A1 and other helpful cytokines to successfully reduce the degree of liver damage and promote liver repair and regeneration.[93] Vascular cell adhesion molecule-1 and very late antigen-4 adhesion molecules help PMSCs to concentrate in injured areas. Through cell-cell interactions and the production of growth factors including hepatocyte growth factor (HGF) and vascular endothelial growth factor (VEGF), PMSCs help to heal the injured liver.[94,95] These stem cells have been shown to induce a notable decrease in fibrosis and also in the serum amounts of transaminases. An enhancement in liver regeneration has been suggested due to initiation in the mechanism of autophagy[96], a decline in apoptosis, repression in the activation of stellate cells[97], and incitement in liver cell proliferation.[98] They are also shown to limit hepatocyte inflammation as well as prevent apoptosis by suppression of TNF-α and IFN-γ. This causes the rejuvenation of hepatocytes and release of HGF, VEGF, IL-6, and PAF (Platelet-activating factor).[75,99] No indications regarding the formation of hepatocytes from cell transplants have been revealed in the rat model exhibiting CC14-

induced fibrosis[98], however, engraftment of PMSCs along with human albumin and α-fetoprotein expression has been outlined.[97,100,101]

Bone and Cartilage Diseases

Although bones have a physiological capability of regeneration and remodeling as evident from the healing of bones in case of any damage, there are some physiological scenarios when the physical regeneration of bone is compromised and insufficient. Osteoporosis is one of the conditions where compromised regeneration is seen. Deployment of PMSCs as a therapy for bone disorders exploits their multipotency. PMSCs possess the potential to differentiate and transform into osteocytes. They are *in vivo* potential of forming bones have been acknowledged by many studies thereby making the therapy an interesting candidate to facilitate bone repair. In a study, PMSCs injected in mice models with femur defects were able to fully recover the damage without any signs of fibrosis. They were injected in combination with nano biphasic calcium phosphate ceramics.[102] Another study used hydroxyapatite/tricalcium phosphate particles to form bones at the implants in a severe immune deficiency (SCID) mice model.[103] The multipotency and ability to differentiate into musculoskeletal lineages makes PMSC an alternative and interesting option for osteoarthritis (OA), a degenerative disease of cartilages in the joints. Many studies have been conducted supporting the regenerative hypothesis using PMSCs. Rat models with osteochondral defects in the knee were found to form hyaline cartilage when administered with the PMSCs grown on silk fibroin.[104] In a study, for covering the femoral defects, PMSCs infused in collagen I gel were subjected to a diseased rat model. Upon testing, the soft-grown tissue was positive for toluidine blue which indicated the differentiation of transplanted cells.[105] Proliferative and osteogenic effects of andrographolide (AP) on human-PMSCs have also been studied. It also improved osteogenic differentiation by increasing the expression of osteoblast-specific mRNA. Furthermore, it can aid bone development by increasing the amount of the bone structural protein osteocalcin in osteoblastic cells.[106]

Intestinal Diseases

Crohn's disease, as well as ulcerative colitis, are chronic illnesses triggered due to prolonged swelling of the epithelial cells in the intestinal tract leading to tissue damage all over the gastrointestinal system. These conditions are thought to be a consequence of an aberrant immunological response to intra-luminal antigens among susceptible individuals. Many other specific gene mutations of the nucleotide-binding oligomerization domain 2 (NOD2) are involved in the progression of this disorder.[107] These two illnesses have significant effects on well-being and presently no cure is available. In addition, most sufferers are not receptive to early detection and treatment. Injecting a conditioned medium of PMSCs intraperitoneally improved clinical characteristics in a mouse paradigm suffering from colitis instigated by dextran sulfate sodium.[108] These stem cells when injected intraperitoneally also inhibited tissue necrosis and reduced mouse fatalities. These positive effects seemed to be significantly higher while using NOD2-activated stem cells.[82]

A study showed the safety of intravenous infusion of these stem cells (PDA001) in individuals suffering from intermediate to serious level Crohn’s disorder, and subsequently, certain recovery rates of the illness were also noted.[109] Similarly, in another monitored clinical experiment, the patient’s situation enhanced dramatically on intravenously injecting these stem cells. This led to a considerable decrease in steroid concentrations. Furthermore, many patients having anal fistula exhibited significant progress.[110]

Clinical trials of placenta-derived mesenchymal stem cells for treating various diseases

Clinical trials are human research studies that are used to assess the effectiveness of pharmacological, surgical, or cognitive interventions. PMSCs have also undergone clinical trials for the treatment of certain diseases. For instance, a Phase 1 study has been conducted for determining the safety of intramuscular dosage of PLX-PAD (placenta-derived MSC) for treating critical limb ischemia. Similarly, recently another phase 1 and 2 clinical trials are being conducted for assessing the potential of cryopreserved PMSCs against deadly SARS-COV-2. Many such registered clinical levels are mentioned in Table 1.

Table 1: List of registered clinical trials of placenta-derived mesenchymal stem cells for treating various diseases.[111]

NCT Number	Purpose of Study	Disease	Phase
NCT00919958 (YOS-2009)	Determining the safety of PLX-PAD (placenta-derived MSC) single dose, intramuscular injection for the treatment of CLI	Critical limb ischemia	Phase 1
NCT04464213 (YOS-2020)	Evaluate the efficacy and safety of human PMSCs on diabetic foot ulcer	Diabetic foot ulcer	Phase 1
NCT04464213 (YOS- 2020)	Study the pharmacokinetics and tolerance of hPMSC on diabetic foot ulcer	Diabetic foot ulcer	Phase 1
NCT01420432 (YOS-2011)	Safety and efficacy study of umbilical cord/PMSCs to treat ankylosing spondylitis	Ankylosing spondylitis (inflammatory arthritis)	Phase 1
NCT01413035 (2011)	Evaluate the safety and efficacy of hPMSCs at a dose of 1.0E+6 MSC/kg for type 2 diabetes	Type 2 diabetes	Phase 1 and 2
NCT01413035 (2011)	Evaluate the safety and efficacy of MSCs derived from the human placenta at a dose of 1.0E+6 MSC/kg for type 2 diabetes	Type 2 diabetes	Phase 1 and 2
NCT04652908 (2020)	Study the efficacy of treatment of human placenta-derived MSCs seeded on a commercially available dural graft extracellular matrix against myelomeningocele	Myelomeningocele	Phase 1 and 2 (ongoing)
NCT04461925 (2020)	Assessing infusions of cryopreserved allogeneic multipotent mesenchymal stem cells of the placenta and umbilical cord for COVID-19 patients with acute respiratory distress syndrome	SARS-CoV-2	Phase 1 and 2 (ongoing)
NCT01129739 (2010)	Evaluate safety and efficacy of hPMSCs at a dose of 1.0E+6 MSC/kg on subjects for refractory anemia (RA) and refractory anemia with ring sideroblast (RARS) of myelodysplastic syndromes (MDS)	Myelodysplastic syndromes (a type of cancer)	Phase 2

Future Prospects and Concluding Remarks

Placenta derived mesenchymal stem cells serve as potential sources for use in stem cell therapy in humans. Cell therapy employing these stem cells is primarily based on three essential properties they possess, i.e., their innate remedial ability or the release of paracrine factors, their ability to homing and engraftment and immune regulation capacity. Conversely, the medical utilization of PMSCs is still in its infant stages and many trials are still under research. Many trails involving cell therapy have been conducted using autologous stem cells. However, there are many drawbacks to using a patient's own cells. Firstly, there is a time constraint, as the growth and quality control of autologous cells can take some months. In addition, cells may exhibit lesser efficacy owing to inherent elderly factors, and these PMSCs pose some attributes which can make autologous gene therapy virtually impossible, as observed in the case of older patients as well as those with particular systemic chronic conditions such as diabetes. Allogeneic mesenchymal stem cells, on the other hand, have the ability to be produced in high yields quickly so that they can be easily accessible and delivered instantly. They can be procured under more systematic and explicitly tested parameters and are likely to minimize costs. Till now, published reports on the efficacy of therapy using PMSCs suggest that these cells are safe and thus they are already 'off-the-shelf' products. While many clinical studies are underway or provide no published findings, there are a few positive reports on effectiveness of this treatment. Regenerative nanomedicine is indeed a propitious field which has provided very satisfactory outcomes at pre-clinical levels. Therapies of many illnesses can profit with the utilization of scaffolds, which provide a three-dimensional framework to facilitate cells, promote its adherence and development, thereby improve engraftment and subsequently the therapeutic outcomes. Apart from the use of cells as transporters of nanomaterials to transport medications within the wounded tissue and, even more, the prospect of stimulus-controlled secretion of medicines appears to be exhilarating. Along the way, further research is required on these stem cells which can potentially give promising results in the field of regenerative medicine.

Acknowledgments

The authors are grateful to the Delhi Technological University, New Delhi, India for providing the opportunity and support to conduct this study.

Conflicts of Interest

The author declares that there are no conflicts of interest relevant to this article.

References

1. Hipp J, Atala A. Sources of Stem Cells for Regenerative Medicine. Stem Cell Rev. 2008 Mar 20;4(1):3–11.
2. Quartu M, Serra MP, Boi M, Ibba V, Melis T, Del Fiacco M. Polysialylated-neural cell adhesion molecule (PSA-NCAM) in the human trigeminal ganglion and brainstem at prenatal and adult ages. BMC Neurosci. 2008 Dec 6;9(1):108.
3. Yen BL, Huang H-I, Chien C-C, Jui H-Y, Ko B-S, Yao M, et al. Isolation of Multipotent Cells from Human Term Placenta. Stem Cells. 2005 Jan;23(1):3–9.
4. Lee OK, Kuo TK, Chen W-M, Lee K-D, Hsieh S-L, Chen T-H. Isolation of multipotent mesenchymal stem cells from umbilical cord blood. Blood. 2004 Mar 1;103(5):1669–75.
5. Napso T, Yong HEJ, Lopez-Tello J, Sferruzzi-Perri AN. The Role of Placental Hormones in Mediating Maternal Adaptations to Support Pregnancy and Lactation. Front Physiol. 2018 Aug 17;9.
6. Zakrzewski W, Dobrzyński M, Szymonowicz M, Rybak Z. Stem cells: past, present, and future. Stem Cell Res Ther. 2019 Dec 26;10(1):68.

7. Macias MI, Grande J, Moreno A, Domínguez I, Bornstein R, Flores AI. Isolation and characterization of true mesenchymal stem cells derived from human term decidua capable of multilineage differentiation into all 3 embryonic layers. Am J Obstet Gynecol. 2010 Nov;203(5):495.e9-495.e23.
8. Anker PS, Scherjon SA, Kleijburg-van der Keur C, de Groot-Swings GMJS, Claas FHJ, Fibbe WE, et al. Isolation of Mesenchymal Stem Cells of Fetal or Maternal Origin from Human Placenta. Stem Cells. 2004 Dec;22(7):1338–45.
9. Fukuchi Y, Nakajima H, Sugiyama D, Hirose I, Kitamura T, Tsuji K. Human Placenta-Derived Cells Have Mesenchymal Stem/Progenitor Cell Potential. Stem Cells. 2004 Sep;22(5):649–58.
10. Yen BL, Chien C-C, Chen Y-C, Chen J-T, Huang J-S, Lee F-K, et al. Placenta-Derived Multipotent Cells Differentiate into Neuronal and Glial Cells In Vitro. Tissue Eng Part A. 2008 Jan;14(1):9–17.
11. Chang C-M, Kao C-L, Chang Y-L, Yang M-J, Chen Y-C, Sung B-L, et al. Placenta-derived multipotent stem cells induced to differentiate into insulin-positive cells. Biochem Biophys Res Commun. 2007 Jun;357(2):414–20.
12. Chien C-C, Yen BL, Lee F-K, Lai T-H, Chen Y-C, Chan S-H, et al. In Vitro Differentiation of Human Placenta-Derived Multipotent Cells into Hepatocyte-Like Cells. Stem Cells. 2006 Jul;24(7):1759–68.
13. Lee JM, Jung J, Lee H-J, Jeong SJ, Cho KJ, Hwang S-G, et al. Comparison of immunomodulatory effects of placenta mesenchymal stem cells with bone marrow and adipose mesenchymal stem cells. Int Immunopharmacol. 2012 Jun;13(2):219–24.
14. Yao P, Zhou L, Zhu L, Zhou B, Yu Q. Mesenchymal Stem Cells: A Potential Therapeutic Strategy for Neurodegenerative Diseases. Eur Neurol. 2020;83(3):235–41.
15. Ns S, Na A. Sources of Mesenchymal Stromal Cells : An Overview. Am Jorunal Pharmacol. 2018;1(1):1–7.
16. Yun Cheng H. The Impact of Mesenchymal Stem Cell Source on Proliferation, Differentiation, Immunomodulation and Therapeutic Efficacy. J Stem Cell Res Ther. 2014;04(10).
17. Pipino C, Shangaris P, Resca E, Zia S, Deprest J, Sebire NJ, et al. Placenta as a reservoir of stem cells: an underutilized resource? Br Med Bull. 2013 Mar 1;105(1):43–68.
18. Kim EY, Lee K-B, Kim MK. The potential of mesenchymal stem cells derived from amniotic membrane and amniotic fluid for neuronal regenerative therapy. BMB Rep. 2014 Mar 31;47(3):135–40.
19. Siddesh SE, Gowda DM, Jain R, Gulati A, Patil GS, Anudeep TC, et al. Placenta-derived mesenchymal stem cells (P-MSCs) for COVID-19 pneumonia-a regenerative dogma. Stem Cell Investig. 2021 Feb;8:3–3.
20. Woods L, Perez-Garcia V, Hemberger M. Regulation of Placental Development and Its Impact on Fetal Growth—New Insights From Mouse Models. Front Endocrinol (Lausanne). 2018 Sep 27;9.
21. Abbasi-Kangevari M, Ghamari S-H, Safaeinejad F, Bahrami S, Niknejad H. Potential Therapeutic Features of Human Amniotic Mesenchymal Stem Cells in Multiple Sclerosis: Immunomodulation, Inflammation Suppression, Angiogenesis Promotion, Oxidative Stress Inhibition, Neurogenesis Induction, MMPs Regulation, and Remyelination Stim. Front Immunol. 2019 Feb 20;10.
22. Siddesh SE, Gowda DM, Jain R, Gulati A, Patil GS, Anudeep TC, et al. Placenta-derived mesenchymal stem cells (P-MSCs) for COVID-19 pneumonia-a regenerative dogma. Stem Cell Investig. 2021 Jul 9;8.
23. Toda A, Okabe M, Yoshida T, Nikaido T. The Potential of Amniotic Membrane/Amnion-Derived Cells for Regeneration of Various Tissues. J Pharmacol Sci. 2007;105(3):215–28.
24. Ambrósio CE, Orlandin JR, Oliveira VC, Motta LCB, Pinto PAF, Pereira VM, et al. Potential application of aminiotic stem cells in veterinary medicine. Anim Reprod. 2020 May 22;16(1):24–30.
25. Portmann-Lanz CB, Schoeberlein A, Huber A, Sager R, Malek A, Holzgreve W, et al. Placental mesenchymal stem cells as potential autologous graft for pre- and perinatal neuroregeneration. Am J Obstet Gynecol. 2006 Mar;194(3):664–73.
26. Ventura Ferreira MS, Bienert M, Müller K, Rath B, Goecke T, Opländer C, et al. Comprehensive characterization of chorionic villi-derived mesenchymal stromal cells from human placenta. Stem Cell Res Ther. 2018 Dec 5;9(1):28.
27. L. PK, Kandoi S, Misra R, S. V, K. R, Verma RS. The mesenchymal stem cell secretome: A new paradigm towards cell-free therapeutic mode in regenerative medicine. Cytokine Growth Factor Rev. 2019

Apr;46:1–9.
28. Bailo M, Soncini M, Vertua E, Signoroni PB, Sanzone S, Lombardi G, et al. Engraftment Potential of Human Amnion and Chorion Cells Derived from Term Placenta. Transplantation. 2004 Nov 27;78(10):1439–48.
29. Vegh I, Grau M, Gracia M, Grande J, de la Torre P, Flores AI. Decidua mesenchymal stem cells migrated toward mammary tumors in vitro and in vivo affecting tumor growth and tumor development. Cancer Gene Ther. 2013 Jan 5;20(1):8–16.
30. Pogozhykh O, Prokopyuk V, Figueiredo C, Pogozhykh D. Placenta and Placental Derivatives in Regenerative Therapies: Experimental Studies, History, and Prospects. Stem Cells Int. 2018;2018:1–14.
31. Gnecchi M, Zhang Z, Ni A, Dzau VJ. Paracrine Mechanisms in Adult Stem Cell Signaling and Therapy. Circ Res. 2008 Nov 21;103(11):1204–19.
32. Taglauer ES, Adams Waldorf KM, Petroff MG. The hidden maternal-fetal interface: events involving the lymphoid organs in maternal-fetal tolerance. Int J Dev Biol. 2010;54(2–3):421–30.
33. Blaschitz A, Hutter H, Dohr G. HLA Class I protein expression in the human placenta. Early Pregnancy. 2001 Jan;5(1):67–9.
34. Xu X, Zhou Y, Wei H. Roles of HLA-G in the Maternal-Fetal Immune Microenvironment. Front Immunol. 2020 Oct 22;11.
35. Bowen JM, Chamley L, Mitchell MD, Keelan JA. Cytokines of the placenta and extra-placental membranes: biosynthesis, secretion and roles in establishment of pregnancy in women. Placenta. 2002 Apr;23(4):239–56.
36. McCracken SA, Hadfield K, Rahimi Z, Gallery ED, Morris JM. NF-κB-regulated suppression of T-bet in T cells represses Th1 immune responses in pregnancy. Eur J Immunol. 2007 May;37(5):1386–96.
37. Heikkinen J, Möttönen M, Alanen A, Lassila O. Phenotypic characterization of regulatory T cells in the human decidua. Clin Exp Immunol. 2004 May;136(2):373–8.
38. Abumaree MH, Al Jumah MA, Kalionis B, Jawdat D, Al Khaldi A, Abomaray FM, et al. Human Placental Mesenchymal Stem Cells (pMSCs) Play a Role as Immune Suppressive Cells by Shifting Macrophage Differentiation from Inflammatory M1 to Anti-inflammatory M2 Macrophages. Stem Cell Rev Reports. 2013 Oct 28;9(5):620–41.
39. Yust-Katz S, Fisher-Shoval Y, Barhum Y, Ben-Zur T, Barzilay R, Lev N, et al. Placental mesenchymal stromal cells induced into neurotrophic factor-producing cells protect neuronal cells from hypoxia and oxidative stress. Cytotherapy. 2012 Jan;14(1):45–55.
40. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. N Engl J Med. 2020 Mar 26 [cited 2021 Jul 4];382(13):1199–207.
41. Madabhavi I, Sarkar M, Kadakol N. COVID-19. A review. Monaldi Arch Chest Dis. 2020 May 14;90(2).
42. Sun P, Lu X, Xu C, Sun W, Pan B. Understanding of COVID-19 based on current evidence. J Med Virol. 2020 Jun 5;92(6):548–51.
43. Luthra R, Kaur S, Bhandari K. Applications of CRISPR as a potential therapeutic. Life Sci. 2021 Nov;284:119908.
44. Iwata-Yoshikawa N, Okamura T, Shimizu Y, Hasegawa H, Takeda M, Nagata N. TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. Gallagher T, editor. J Virol. 2019 Mar 15;93(6).
45. Leng Z, Zhu R, Hou W, Feng Y, Yang Y, Han Q, Shan G, Meng F, Du D, Wang S, Fan J, Wang W, Deng L, Shi H, Li H, Hu Z, Zhang F, Gao J, Liu H, Li X, Zhao Y, Yin K, He X, Gao Z, Wang Y, Yang B, Jin R, Stambler I, Lim LW, Su H, Moskalev A, Cano A, Chakrabarti S, Min KJ, Ellison-Hughes G, Caruso C, Jin K, Zhao RC. Transplantation of ACE2- Mesenchymal Stem Cells Improves the Outcome of Patients with COVID-19 Pneumonia. Aging Dis. 2020 Mar 9;11(2):216-228.
46. Finegold JA, Asaria P, Francis DP. Mortality from ischaemic heart disease by country, region, and age: Statistics from World Health Organisation and United Nations. Int J Cardiol. 2013 Sep;168(2):934–45.
47. Faiella W, Atoui R. Therapeutic use of stem cells for cardiovascular disease. Clin Transl Med. 2016 Dec 18;5(1).

48. Zhao P, Ise H, Hongo M, Ota M, Konishi I, Nikaido T. Human Amniotic Mesenchymal Cells Have Some Characteristics of Cardiomyocytes. Transplantation. 2005 Mar 15;79(5):528–35.
49. Okamoto K, Miyoshi S, Toyoda M, Hida N, Ikegami Y, Makino H, et al. 'Working' cardiomyocytes exhibiting plateau action potentials from human placenta-derived extraembryonic mesodermal cells. Exp Cell Res. 2007 Jul;313(12):2550–62.
50. Li L, Jaiswal PK, Makhoul G, Jurakhan R, Selvasandran K, Ridwan K, et al. Hypoxia modulates cell migration and proliferation in placenta-derived mesenchymal stem cells. J Thorac Cardiovasc Surg. 2017 Aug;154(2):543-552.e3.
51. Lpez Y, Lutjemeier B, Seshareddy K, M. Trevino E, Sue Hageman K, I. Musch T, et al. Wharton's Jelly or Bone Marrow Mesenchymal Stromal Cells Improve Cardiac Function Following Myocardial Infarction for More Than 32 Weeks in a Rat Model: A Preliminary Report. Curr Stem Cell Res Ther. 2013 Feb 1;8(1):46–59.
52. Simioniuc A, Campan M, Lionetti V, Marinelli M, Aquaro GD, Cavallini C, et al. Placental stem cells pre-treated with a hyaluronan mixed ester of butyric and retinoic acid to cure infarcted pig hearts: a multimodal study. Cardiovasc Res. 2011 Jun 1;90(3):546–56.
53. Danieli P, Malpasso G, Ciuffreda MC, Cervio E, Calvillo L, Copes F, et al. Conditioned Medium From Human Amniotic Mesenchymal Stromal Cells Limits Infarct Size and Enhances Angiogenesis. Stem Cells Transl Med. 2015 May;4(5):448–58.
54. Gao LR, Chen Y, Zhang NK, Yang XL, Liu HL, Wang ZG, et al. Intracoronary infusion of Wharton's jelly-derived mesenchymal stem cells in acute myocardial infarction: double-blind, randomized controlled trial. BMC Med. 2015 Dec 10;13(1):162.
55. Bartolucci J, Verdugo FJ, González PL, Larrea RE, Abarzua E, Goset C, et al. Safety and Efficacy of the Intravenous Infusion of Umbilical Cord Mesenchymal Stem Cells in Patients With Heart Failure. Circ Res. 2017 Oct 27;121(10):1192–204.
56. Yu S, You X, Liang H, Li Y, Fu Y, Zhang X, Hu X, An J, Xu Y, Li F. First trimester placental mesenchymal stem cells improve cardiac function of rat after myocardial infarction via enhanced neovascularization. Heliyon. 2021 Feb 1;7(1):e06120.
57. Sprengers RW, Teraa M, Moll FL, de Wit GA, van der Graaf Y, Verhaar MC. Quality of life in patients with no-option critical limb ischemia underlines the need for new effective treatment. J Vasc Surg. 2010 Oct;52(4):843-849.e1.
58. Xie B, Luo H, Zhang Y, Wang Q, Zhou C, Xu D. Autologous Stem Cell Therapy in Critical Limb Ischemia: A Meta-Analysis of Randomized Controlled Trials. Stem Cells Int. 2018;2018:1–12.
59. Gu Y, Rampin A, Alvino V V., Spinetti G, Madeddu P. Cell Therapy for Critical Limb Ischemia: Advantages, Limitations, and New Perspectives for Treatment of Patients with Critical Diabetic Vasculopathy. Curr Diab Rep. 2021 Mar 2;21(3):11.
60. Beltrán-Camacho L, Rojas-Torres M, Durán-Ruiz MC. Current Status of Angiogenic Cell Therapy and Related Strategies Applied in Critical Limb Ischemia. Int J Mol Sci. 2021 Feb 26;22(5):2335.
61. Prather WR, Toren A, Meiron M, Ofir R, Tschope C, Horwitz EM. The role of placental-derived adherent stromal cell (PLX-PAD) in the treatment of critical limb ischemia. Cytotherapy. 2009 Jan;11(4):427–34.
62. Xie N, Li Z, Adesanya TM, Guo W, Liu Y, Fu M, et al. Transplantation of placenta-derived mesenchymal stem cells enhances angiogenesis after ischemic limb injury in mice. J Cell Mol Med. 2016 Jan 18;20(1):29–37.
63. Liang L, Li Z, Ma T, Han Z, Du W, Geng J, et al. Transplantation of Human Placenta-Derived Mesenchymal Stem Cells Alleviates Critical Limb Ischemia in Diabetic Nude Rats. Cell Transplant. 2017 Jan 1;26(1):45–61.
64. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart Disease and Stroke Statistics-2016 Update. Circulation. 2016 Jan 26;133(4).
65. Zahavi-Goldstein E, Blumenfeld M, Fuchs-Telem D, Pinzur L, Rubin S, Aberman Z, et al. Placenta-derived PLX-PAD mesenchymal-like stromal cells are efficacious in rescuing blood flow in hind limb ischemia

mouse model by a dose- and site-dependent mechanism of action. Cytotherapy. 2017 Dec;19(12):1438–46.

66. Barzegar M, Wang Y, Eshaq RS, Yun JW, Boyer CJ, Cananzi SG, et al. Human placental mesenchymal stem cells improve stroke outcomes via extracellular vesicles-mediated preservation of cerebral blood flow. EBioMedicine. 2021 Jan;63:103161.
67. Luthra R, Datta S, Roy A. Role of Different Peptides for Cancer Immunotherapy. Int J Pept Res Ther. 2021 Sep 23.
68. Moodley Y, Vaghjiani V, Chan J, Baltic S, Ryan M, Tchongue J, et al. Anti-Inflammatory Effects of Adult Stem Cells in Sustained Lung Injury: A Comparative Study. Mezey E, editor. PLoS One. 2013 Aug 1;8(8):e69299.
69. Spaeth E, Klopp A, Dembinski J, Andreeff M, Marini F. Inflammation and tumor microenvironments: defining the migratory itinerary of mesenchymal stem cells. Gene Ther. 2008 May 10;15(10):730–8.
70. Aboody KS, Brown A, Rainov NG, Bower KA, Liu S, Yang W, et al. Neural stem cells display extensive tropism for pathology in adult brain: Evidence from intracranial gliomas. Proc Natl Acad Sci. 2000 Nov 7;97(23):12846–51.
71. Auffinger B, Morshed R, Tobias A, Cheng Y, Ahmed AU, Lesniak MS. Drug-Loaded Nanoparticle Systems And Adult Stem Cells: A Potential Marriage For The Treatment Of Malignant Glioma? Oncotarget. 2013 Mar 31;4(3):378–96.
72. Mooney R, Roma L, Zhao D, Van Haute D, Garcia E, Kim SU, et al. Neural Stem Cell-Mediated Intratumoral Delivery of Gold Nanorods Improves Photothermal Therapy. ACS Nano. 2014 Dec 23;8(12):12450–60.
73. Li L, Guan Y, Liu H, Hao N, Liu T, Meng X, et al. Silica Nanorattle–Doxorubicin-Anchored Mesenchymal Stem Cells for Tumor-Tropic Therapy. ACS Nano. 2011 Sep 27;5(9):7462–70.
74. Shen C-J, Chan T-F, Chen C-C, Hsu Y-C, Long C-Y, Lai C-S. Human umbilical cord matrix-derived stem cells expressing interferon-β gene inhibit breast cancer cells via apoptosis. Oncotarget. 2016 Jun 7;7(23):34172-9.
75. Zhang G-Z, Sun H-C, Zheng L-B, Guo J-B, Zhang X-L. In vivo hepatic differentiation potential of human umbilical cord-derived mesenchymal stem cells: Therapeutic effect on liver fibrosis/cirrhosis. World J Gastroenterol. 2017 Dec 14;23(46):8152–68.
76. Ganta C, Chiyo D, Ayuzawa R, Rachakatla R, Pyle M, Andrews G, et al. Rat Umbilical Cord Stem Cells Completely Abolish Rat Mammary Carcinomas with No Evidence of Metastasis or Recurrence 100 Days Post–Tumor Cell Inoculation. Cancer Res. 2009 Mar 1;69(5):1815–20.
77. Stuckey DW, Shah K. TRAIL on trial: preclinical advances in cancer therapy. Trends Mol Med. 2013 Nov;19(11):685–94.
78. De la Torre P, Paris JL, Fernández-de la Torre M, Vallet-Regí M, Flores AI. Endostatin Genetically Engineered Placental Mesenchymal Stromal Cells Carrying Doxorubicin-Loaded Mesoporous Silica Nanoparticles for Combined Chemo- and Antiangiogenic Therapy. Pharmaceutics. 2021 Feb 10;13(2):244.
79. Luthra R, Roy A. Role of medicinal plants against neurodegenerative diseases. Curr Pharm Biotechnol. 2021 Feb 12;22.
80. Park S, Kim E, Koh S-E, Maeng S, Lee W, Lim J, et al. Dopaminergic differentiation of neural progenitors derived from placental mesenchymal stem cells in the brains of Parkinson's disease model rats and alleviation of asymmetric rotational behavior. Brain Res. 2012 Jul;1466:158–66.
81. Shin JY, Lee PH. Mesenchymal stem cells modulate misfolded α-synuclein in parkinsonian disorders: A multitarget disease-modifying strategy. Stem Cell Res. 2020 Aug;47:101908.
82. Kim H, Shin T, Lee B, Yu K, Seo Y, Lee S, et al. Human Umbilical Cord Blood Mesenchymal Stem Cells Reduce Colitis in Mice by Activating NOD2 Signaling to COX2. Gastroenterology. 2013 Dec;145(6):1392-1403.e8.
83. Jiao H, Shi K, Zhang W, Yang L, Yang L, Guan F, et al. Therapeutic potential of human amniotic membrane-derived mesenchymal stem cells in APP transgenic mice. Oncol Lett. 2016 Sep;12(3):1877–83.

84. Garbuzova-Davis S, Rodrigues MCO, Mirtyl S, Turner S, Mitha S, Sodhi J, et al. Multiple Intravenous Administrations of Human Umbilical Cord Blood Cells Benefit in a Mouse Model of ALS. Mosley RL, editor. PLoS One. 2012 Feb 3;7(2):e31254.
85. Fisher-Shoval Y, Barhum Y, Sadan O, Yust-Katz S, Ben-Zur T, Lev N, et al. Transplantation of Placenta-Derived Mesenchymal Stem Cells in the EAE Mouse Model of MS. J Mol Neurosci. 2012 Sep 26;48(1):176–84.
86. Bravo B, Gallego MI, Flores AI, Bornstein R, Puente-Bedia A, Hernández J, et al. Restrained Th17 response and myeloid cell infiltration into the central nervous system by human decidua-derived mesenchymal stem cells during experimental autoimmune encephalomyelitis. Stem Cell Res Ther. 2016 Dec 17;7(1):43.
87. Jiang H, Zhang Y, Tian K, Wang B, Han S. Amelioration of experimental autoimmune encephalomyelitis through transplantation of placental derived mesenchymal stem cells. Sci Rep. 2017 Mar 10;7(1):41837.
88. Li J-F, Zhang D-J, Geng T, Chen L, Huang H, Yin H-L, et al. The Potential of Human Umbilical Cord-Derived Mesenchymal Stem Cells as a Novel Cellular Therapy for Multiple Sclerosis. Cell Transplant. 2014 Jan 1;23(1_suppl):113–22.
89. Lublin FD, Bowen JD, Huddlestone J, Kremenchutzky M, Carpenter A, Corboy JR, et al. Human placenta-derived cells (PDA-001) for the treatment of adults with multiple sclerosis: A randomized, placebo-controlled, multiple-dose study. Mult Scler Relat Disord. 2014 Nov;3(6):696–704.
90. Schuppan D, Afdhal NH. Liver cirrhosis. Lancet. 2008 Mar;371(9615):838–51.
91. Williams R. Sherlock's disease of the liver and biliary systems. Clin Med (Northfield Il). 2011 Oct 1;11(5):506–506.
92. Bornstein R, Macias MI, de la Torre P, Grande J, Flores AI. Human decidua-derived mesenchymal stromal cells differentiate into hepatic-like cells and form functional three-dimensional structures. Cytotherapy. 2012 Sep;14(10):1182–92.
93. Zagoura D, Trohatou O, Makridakis M, Kollia A, Kokla N, Mokou M, et al. Functional secretome analysis reveals Annexin-A1 as important paracrine factor derived from fetal mesenchymal stem cells in hepatic regeneration. EBioMedicine. 2019 Jul;45:542–52.
94. Eom YW, Shim KY, Baik SK. Mesenchymal stem cell therapy for liver fibrosis. Korean J Intern Med. 2015 Aug 27;30(5):580–9.
95. Luan Y, Kong X, Feng Y. Mesenchymal stem cells therapy for acute liver failure: Recent advances and future perspectives. Liver Res. 2021 Jun;5(2):53–61.
96. Jung J, Choi JH, Lee Y, Park J-W, Oh I-H, Hwang S-G, et al. Human Placenta-Derived Mesenchymal Stem Cells Promote Hepatic Regeneration in CCl 4 -Injured Rat Liver Model via Increased Autophagic Mechanism. Stem Cells. 2013 Aug;31(8):1584–96.
97. Zhang D, Jiang M, Miao D. Transplanted Human Amniotic Membrane-Derived Mesenchymal Stem Cells Ameliorate Carbon Tetrachloride-Induced Liver Cirrhosis in Mouse. Aziz S, editor. PLoS One. 2011 Feb 4;6(2):e16789.
98. Tsai P-C, Fu T-W, Chen Y-MA, Ko T-L, Chen T-H, Shih Y-H, et al. The therapeutic potential of human umbilical mesenchymal stem cells from Wharton's jelly in the treatment of rat liver fibrosis. Liver Transplant. 2009 May;15(5):484–95.
99. Fiore EJ, Bayo JM, Garcia MG, Malvicini M, Lloyd R, Piccioni F, et al. Mesenchymal Stromal Cells Engineered to Produce IGF-I by Recombinant Adenovirus Ameliorate Liver Fibrosis in Mice. Stem Cells Dev. 2015 Mar 15;24(6):791–801.
100. Cao H, Yang J, Yu J, Pan Q, Li J, Zhou P, et al. Therapeutic potential of transplanted placental mesenchymal stem cells in treating Chinese miniature pigs with acute liver failure. BMC Med. 2012 Dec 6;10(1):56.
101. Yu J, Hao G, Wang D, Liu J, Dong X, Sun Y, et al. Therapeutic Effect and Location of GFP-Labeled Placental Mesenchymal Stem Cells on Hepatic Fibrosis in Rats. Stem Cells Int. 2017;2017:1–11.
102. Reddy S, Wasnik S, Guha A, Kumar JM, Sinha A, Singh S. Evaluation of nano-biphasic calcium phosphate ceramics for bone tissue engineering applications: In vitro and preliminary in vivo studies. J Biomater Appl. 2013 Jan 27;27(5):565–75.

103. Kusuma GD, Menicanin D, Gronthos S, Manuelpillai U, Abumaree MH, Pertile MD, et al. Ectopic Bone Formation by Mesenchymal Stem Cells Derived from Human Term Placenta and the Decidua. Matsusaki M, editor. PLoS One. 2015 Oct 20;10(10):e0141246.
104. Li F, Chen Y-Z, Miao Z-N, Zheng S, Jin J. Human Placenta-Derived Mesenchymal Stem Cells with Silk Fibroin Biomaterial in the Repair of Articular Cartilage Defects. Cell Reprogram. 2012 Aug;14(4):334–41.
105. Wei JP, Nawata M, Wakitani S, Kametani K, Ota M, Toda A, et al. Human Amniotic Mesenchymal Cells Differentiate into Chondrocytes. Cloning Stem Cells. 2009 Mar;11(1):19–26.
106. Phunikom N, Boonmuen N, Kheolamai P, Suksen K, Manochantr S, Tantrawatpan C, et al. Andrographolide promotes proliferative and osteogenic potentials of human placenta-derived mesenchymal stem cells through the activation of Wnt/β-catenin signaling. Stem Cell Res Ther. 2021 Dec 14;12(1):241.
107. Hugot J-P, Chamaillard M, Zouali H, Lesage S, Cézard J-P, Belaiche J, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature. 2001 May;411(6837):599–603.
108. Song J, Kang HJ, Hong JS, Kim CJ, Shim J-Y, Lee CW, et al. Umbilical cord-derived mesenchymal stem cell extracts reduce colitis in mice by re-polarizing intestinal macrophages. Sci Rep. 2017 Dec 25;7(1):9412.
109. Mayer L, Pandak WM, Melmed GY, Hanauer SB, Johnson K, Payne D, et al. Safety and Tolerability of Human Placenta-Derived Cells (PDA001) in Treatment-Resistant Crohn's Disease. Inflamm Bowel Dis. 2013 Mar;19(4):754–60.
110. Zhang J, Lv S, Liu X, Song B, Shi L. Umbilical Cord Mesenchymal Stem Cell Treatment for Crohn's Disease: A Randomized Controlled Clinical Trial. Gut Liver. 2018 Jan 15;12(1):73–8.
111. What Are Clinical Trials and Studies? National Institute on Aging. https://www.nia.nih.gov/health/what-are-clinical-trials-and-studies. Accessed on 5 October 2021.

Cite this article

Luthra R. Yadav S. Verma S. Therapeutic applications of mesenchymal cells derived from human placenta: A comprehensive meta-review**.** NeuroPharmac J. 2021; 6(3): 277-294. DOI: 10.37881/1.637

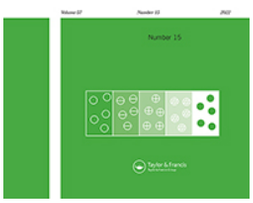
Copyright

© 2021 NeuroPharmac J. This is an open-access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License.

Taylor & Francis
Taylor & Francis Group

SEPARATION SCIENCE AND TECHNOLOGY

Separation Science and Technology



ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lsst20

Thermodynamics, kinetics and isotherm studies on the removal of anionic Azo-dye (Congo red) using synthesized Chitosan/*Moringa oleifera* gum hydrogel composites

Sakshi Nangia, Deeksha Katyal & Sudhir G. Warkar

To cite this article: Sakshi Nangia, Deeksha Katyal & Sudhir G. Warkar (2022): Thermodynamics, kinetics and isotherm studies on the removal of anionic Azo-dye (Congo red) using synthesized Chitosan/*Moringa oleifera* gum hydrogel composites, Separation Science and Technology, DOI: 10.1080/01496395.2022.2104731

To link to this article: https://doi.org/10.1080/01496395.2022.2104731

Published online: 02 Aug 2022.

Submit your article to this journal

Article views: 65

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=lsst20

SEPARATION SCIENCE AND TECHNOLOGY
https://doi.org/10.1080/01496395.2022.2104731

Taylor & Francis
Taylor & Francis Group

Check for updates

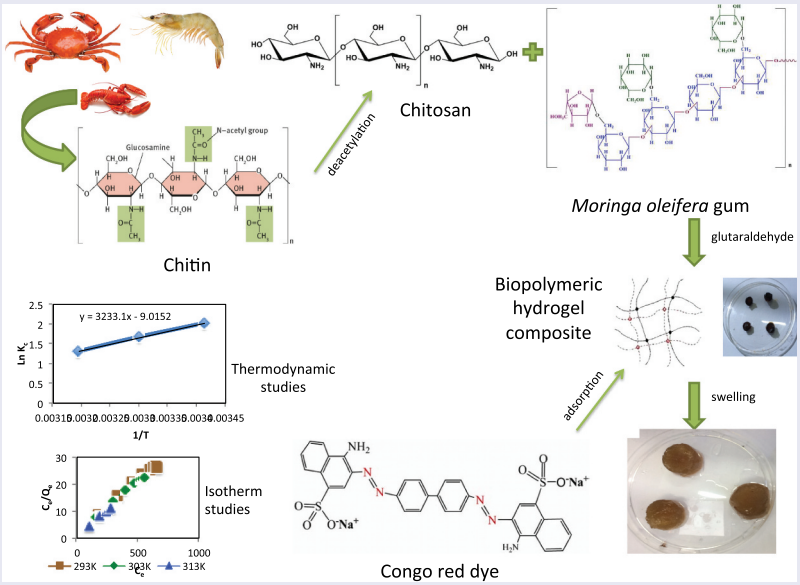
Thermodynamics, kinetics and isotherm studies on the removal of anionic Azo-dye (Congo red) using synthesized Chitosan/*Moringa oleifera* gum hydrogel composites

Sakshi Nangia[a], Deeksha Katyal[a], and Sudhir G. Warkar[b]

[a]University School of Environment Management, Guru Gobind Singh Indraprastha University, New Delhi, India; [b]Department of Applied Chemistry, Delhi Technological University, New Delhi, India

ABSTRACT

Low cost and biocompatible hydrogels were synthesized using chitosan (CS) and *Moringa oleifera* (MO) gum biopolymers through Maillard reaction with glutaraldehyde as crosslinker. These CS/MO hydrogels were characterized using FT-IR spectroscopy and thermo gravimetric analysis (TGA). Equilibrium swelling behavior at different pH, temperature and salt solutions indicated almost equal swelling at all pH values, while higher swelling was observed at high temperature and low salt concentrations. Congo red dye, an azo dye, is potentially carcinogenic and can have different molecular structure at different pH. For its removal, variables including initial dye concentration, temperature, pH and contact time were optimized in order to achieve best adsorption capacity. The kinetics and isotherm studies suggested that the process of Congo red sorption on hydrogels was better fitted through pseudo second-order and Langmuir isotherm model. Maximum adsorption capacity of 50.25 mg/g was obtained at pH 2 as per Langmuir isotherm that represented a monolayer adsorption mechanism. The thermodynamic evaluation of the system suggested chemical nature of adsorption of dye. Since CS/MO hydrogels showed removal of more than 95% at all pH values along with ease of use, they could exhibit a potential of commercial utility in textile sectors for Congo red dye removal if further studies are conducted.



ARTICLE HISTORY
Received 24 January 2022
Accepted 18 July 2022

KEYWORDS
Chitosan; *moringa oleifera* gum; Congo red dye; thermodynamics; isotherms

Introduction

The essential entities of life such as air and water have become dump yards of pollutants leading to varied diseases and health issues in different living organisms. Our bodies are being exposed to numerous chemicals originating through known or unknown sources in our daily lives. Effluents from textile industries containing various dyes have the tendency of increasing the Biochemical

CONTACT Deeksha Katyal deekshakatyal@gmail.com University School of Environment Management, Guru Gobind Singh Indraprastha University, New Delhi, India

© 2022 Taylor & Francis Group, LLC

Oxygen Demand (BOD) of water, inhibit plant growth and may cause cancer and some other skin diseases. Several histological stains and dyes such as Congo red, being used in textile and pharmaceutical sectors are carcinogenic and act like recalcitrant molecules that persist in nature for long durations.[1] Congo red is an anionic azo dye that can be metabolized in the body into benzidene, which causes cancer.[2] As compared to other anionic dyes, Congo red dye can have different molecular structures at different pH of solution, leading to difficulty in attachment to adsorbents and thus difficulty in removal. Thus, there is an urgent need to remove this dye from wastewater because of the serious implications it carries. The conventional treatments for dye removal that involve physico-chemical processes such as advanced oxidation process or alternatively, biological processes such as using biofilms and tertiary treatments such as adsorption through activated carbon have been used extensively over many years.[3] However, all the above-mentioned three processes require high dosage of chemicals and energy, are relatively expensive and lead to complex sludge generation.[4] To effectively address these issues, focus has been shifted to the use of biodegradable, low cost and available sorbents.

Hydrogels are hydrophilic polymers, which have a swelling percentage of more than 100 for water or other aqueous/biological fluids.[5–7] Petroleum-based vinyl monomers such as acrylates are mostly used for the synthesis of hydrogels and continue to find use in commercial applications such as in baby diapers. However, these hydrogels suffer with the challenge of not being ecofriendly and possessing a low degradability. Thus, green polymers like starch, cellulose, chitosan, cashew gum and yeast or derived and modified biopolymers like chitosan-locust bean gum nanocomposites have come out as an alternative solution.[8] These green polymers are biodegradable in nature and are termed as biopolymers.

Chitosan (CS) is an emerging sorbent that is obtained from chitin derived from waste crustacean shells. The sorption ability is due to the availability of $-NH_2$ and –OH groups on its surface.[9] However, the disadvantage associated with using chitosan is that it lacks mechanical strength and is easily disintegrated in acidic environment. Thus, efforts have been made to combine and test it with several synthetic counterparts like N,N'-dimethyl acrylamide[10] and polymethacrylic acid,[11] leading to improvement in its mechanical strength.

Hydrogels of CS with other biopolymers is less common as compared to synthetic polymers, as the synthetic counterparts offer better mechanical strength and swelling ability.[12] Thus, more research directed toward mixing chitosan with other biopolymers is desired to ensure safety for environment while improving properties such as adsorption/coagulation properties and mechanical strength of these gels.[13]

These biopolymers required to augment the properties of CS could be plant-derived extracts, which are low in cost and are abundantly available. A great alternative is the stem of the widely distributed drumstick plant with scientific name *Moringa oleifera* (MO) that has been studied for its gelling nature[14] and has been used as a binding agent.[15] The Indian name of *Moringa oleifera* is "Shobhanjana/saijna," which is known to be an auspicious tree.

Though there are several studies on adsorption of dyes and heavy metals using different chitosan-based composites and hydrogels, our relative focus is essentially on MO gum. Essentially, the studies available in literature detail out adsorption processes using MO leaf/seed extracts.[16,17] CS-MO gum hydrogel is a novel combination which has never been used for removal of Congo red dye from its aqueous solution. Moreover, it has been observed during the experiments conducted that the adsorption capacity of the adsorbent for removal of dye was increased by about 5 times due to the presence of MO gum in the hydrogel. More adsorption capacity indicates availability of more active sites and thus proportional dye uptake for a particular dosage of adsorbent. This also means there is more efficiency of removal of dye. Thus, this study differentiates itself in generating an amalgam of two green polymers namely CS and MO gum to find its designated use in the sorption of anionic dyes. To give a fair idea of novelty, a further review is presented drawing comparisons with the available literature compared with the following studies which are cited in Table 4.

CS has received considerable attention in environmental arenas because of its biocompatibility. It is a promising sorbent with which slight increase in mechanical strength would increase its utility manifold in diverse areas. Ideally, this modification can be achieved by its crosslinking with chemical agents such as glutaraldehyde or genipin or formation of various composites by combining it with other polymers. Some combinations of CS such as DNA-CS hydrogel, Karaya gum-CS sponge hydrogel and terephthaloyl thiourea crosslinked CS hydrogels have been used for adsorption of dyes.[2,21,22] However, there are no studies of dye adsorption through MO gum in combination with CS. Limited studies are available on dye removal through modified MO gum, with hardly one or two authors reporting about its composites for removal of dyes.[23] The removed dye in these limited studies was malachite green and not Congo red dye as used in this study.

Individually, some studies available on MO gum are presented in this section. MO gum-based polyurethane foam with ash as filler has been used for the removal of malachite green dye.[23] Also, MO gum modified using etherification by alkyl halides has also been used for the removal of the same dye malachite green.[24] *Moringa oleifera* gum and bentonite clay combination also finds usage in the synthesis of a composite coagulant which has been used for the removal of heavy metals from aqueous systems for drinking water purposes.[25] These studies were indicative of the potential use of MO gum both as an adsorbent and as a coagulant. Some recent studies were suggestive of utility of MO gum in the synthesis of hydrogel. One such study was the synthesis of hydrogel from MO gum and polyvinyl alcohol which was crosslinked using borax.[26] Another MO gum-based hydrogel was synthesized with poly acrylic acid by radiation-induced crosslinking which was further explored for drug delivery applications.[27]

As is evident, these studies have explored the usage of CS and MO gum individually in the field of drug delivery and removal of potential contaminant. However, none of them have used this unique combination, i.e. CS/MO gum hydrogel for removal of dyes which has been delved upon in our study.

Using the two biopolymers, synthesis of hydrogels was carried out using their different ratios and their crosslinking using glutaraldehyde. It was observed while trying various combinations of CS and MO gum that the presence of MO gum enhances the sorption capacity of CS (which is about 10 mg/g) by about 5 times. The stabilized combination of CS and MO gum was duly characterized using FTIR, TGA, DSC and SEM. The first stage involved assessment of swelling properties of the gel at different pH, temperature and salt concentrations. Swelling studies were essential in order to predict the applicability of the synthesized hydrogel. This was followed by a second stage, which comprised of evaluating the sorption properties of the gel through an anionic dye, in this case Congo red, which was selected keeping in mind the overall cationic nature of the hydrogel. The sorption studies were also conducted at different pH and temperature conditions keeping in mind the large influence these parameters have on the removal of the dye. The third stage involved modeling of isotherms and kinetics, which gave an insight into the thermodynamics of the process, further helping in determining the nature of sorption involved. It was found that this combination had the ability of blending the coagulating nature of MO gum with sorbent properties of CS, which finally resulted in almost complete removal of the dye. It is also important to bring forth that the sorbent has been applied in the form of dry pellets, which can be easily removed as compared to powders. The encouraging results in terms of removal at different pH reiterated that this unique combination could prove its mettle in the areas of textile and histological sectors.

Materials and methods

Materials

Chitosan (428850500, Acros Organics) also known as deacetylated chitin or Poliglusam having molecular weight 600,000–800,000 (because of which the viscosity lies in the range of 50–200 mPa.s) and *Moringa oleifera* gum known as Shobhanjna/Saijna gum (MOGUM101-11, Agasthiya Herbs and Spices, Karnataka, India) were procured and used as such. Glutaraldehyde (also known as Cidex or Glutaral in India) 25% solution (83489, Thomas Baker) having molecular weight 100.12 was diluted to 1% before use. Acetic acid Glacial (21057, Fisher Scientific) were procured and used as such. All the reagents were prepared using double distilled water. Wherever required, pH was adjusted using 0.1 M HCl and 0.1 M NaOH till the desired value was achieved.

Methods

Synthesis of CS/MO hydrogel composites

Different dosages of CS were mixed in 10 mL of 1% (v/v) acetic acid.[28] MO gum dispersed in 10 mL distilled water was added to chitosan solution and mixed for 1 hour through a magnetic stirrer at around 480 rpm (at 40°C) in 100 mL beaker, which were then transferred to test tubes. Appropriate volume of crosslinker glutaraldehyde (1% concentration) was added to the above biopolymeric solution. The total content was kept as 50 mL by adding distilled water. The contents were then transferred to test tubes of 50 mL capacity and were thoroughly mixed for about 5 minutes for the formation of hydrogels. The test tubes were kept at room temperature, and hydrogels were formed within an hour. The tubes were then broken, and hydrogels were removed and cut in the form of equal sized cylindrical disks. Cutting the hydrogels properly in uniform dimensions is important for further determining the swelling kinetics. The disks were immersed in distilled water, rinsed, swollen for 1 day and kept for drying in oven at 35°C. The different combinations prepared for getting good hydrogel in terms of texture, strength and handling are described in Table 1.

Table 1. Compositions of the hydrogel prepared.

Notation of hydrogels synthesized	Chitosan (g)	*Moringa oleifera* gum (g)	Glutaraldehyde (1%) (mL)	State of hydrogel formation
$CS_{0.75}MO_{0.25}Gl_{0.4}$/ CS/MO	0.75	0.25	0.4	Hydrogel formed
$CS_{0.5}MO_{0.5}Gl_{0.4}$	0.5	0.5	0.4	Hydrogel formed but texture not smooth
$CS_{0.25}MO_{0.75}Gl_{0.4}$	0.25	0.75	0.4	Flaccid gel
$CS_{0.75}MO_{0.25}Gl_{0.2}$	0.75	0.25	0.2	Flaccid gel

Characterization of Hydrogel

(a) FT-IR analysis

The FTIR data for CS, MO gum, CS crosslinked with glutaraldehyde hydrogels and CS-cl-MO hydrogels were generated using KBr discs of the powdered samples. Agilent Technologies FTIR Spectrophotometer was used for this purpose in the region 4000–650 cm^{-1} at a resolution of 4 cm^{-1}.

(b) Thermo-gravimetric Analysis (TGA)

The stability of the hydrogel at different temperatures was determined using TGA (Model Q-500, TA Instruments, USA) under nitrogen atmosphere at rate of heating of 10°C/min. The analysis was performed using dry hydrogel sample within a temperature range of 30–500°C under dry nitrogen atmosphere. Dry hydrogel sample was obtained by drying the hydrogel at 40°C in oven till constant weight was obtained.

(c) Scanning Electron Microscopy (SEM)

Since an assessment of the gel's surface roughness/ smoothness and porous structure is essential, surface morphology studies were conducted for hydrogel before and after Congo red dye adsorption using SEM (Model JEOL JSM).

Swelling studies

Swelling Ratio (SR)

Determination of swelling ratio is very important for hydrogels as it helps in the determination of the extent of absorption of solvent and their further applicability for the release of drugs or other compounds.[29] If swelling is high, the gel could be used for applications such as release of fertilizers/ holding of water but would be unsuitable for column studies of adsorption. Swelling ratio of the hydrogel composite was assessed using gravimetric measurement. The dried disc-shaped xerogel sample was placed in a 100 mL beaker filled with 50 mL distilled water at 25°C. The swelling of hydrogel started quickly, and the trend was checked on hourly basis. For this, the swollen hydrogel was regularly withdrawn from the beaker at different time intervals and wiped with a blotting sheet. Then it was weighed and kept back in the beaker.[30] The equation used was:

$$SR = \frac{Weight\ of\ swollen\ sample}{Weight\ of\ dry\ sample} \tag{1}$$

(b) Equilibrium Swelling Ratio (ES%)

This parameter was calculated in order to find the maximum/ equilibrium swelling percentage of the hydrogel. The ES% of hydrogel was assessed by keeping dried disc-shaped hydrogel in distilled water at different pH, temperatures and salt solutions by using the following equation[30]:

$$ES\% = \frac{W_{sw} - W_{dr}}{W_{dr}} \times 100 \tag{2}$$

where W_{dr} is the weight of the dry hydrogel and W_{sw} is the weight of the swollen hydrogel.

Adsorption kinetics

In order to find out the adsorption potential of adsorbent, or the amount of dye, which gets adsorbed over the surface of adsorbent, the following mass balance equation was used[31]:

$$Q_t = \frac{(C_o - C_e)V}{m} \tag{3}$$

Q_t is the quantity of adsorbed dye at a particular time (mg/g)

C_o is the initial dye concentration (mg/L)

C_e is the dye concentration in the filtered supernatant at any measured time interval (mg/L)

V is dye volume taken in conical flask (L) and m is the dry weight of hydrogel used (g)

In order to find out the % Adsorption or dye removal achieved, the following equation has been used:

$$\%\ Adsorption = \frac{(C_o - C_e) \times 100}{C_o} \tag{4}$$

The mechanism of adsorption or how the solute interacts with the adsorbent could be found out by examining the kinetics and thermodynamics.[32]

Several kinetic models have been used for adsorption systems in order to describe the reaction order based on solution concentration.[33] Pseudo first order and Pseudo second order are some of these models based on solution concentration where the adsorbent and solvent concentration remains constant unlike the first- and second-order models.

Pseudo first order could be determined using the following equation:

$$Log(Q_e - Q_t) = \log Q_e - \frac{k_{Pf}}{2.303} t \tag{5}$$

where k_{Pf} is the pseudo first-order rate constant. In case of linear relationship between Log (Q_e-Q_t) and t, the dye adsorption is governed by Pseudo first-order kinetics.

Pseudo second order could be determined using the following equation:

$$\frac{t}{Q_t} = \frac{1}{k_2 Q_{e2}} + \frac{1}{Q_{e2}} t \tag{6}$$

A plot between t/ Q_t and t could be plotted to check for linearity and find the values of slope and intercept for k_2 and Q_{e2}.

Another equation used based on adsorption capacity that has been used to describe adsorption of gases/pollutants is Elovich model/equation.[34,35]

Elovich model

This model is used to predict the mass and energies activating/deactivating the adsorption system.[36] As per this model, the rate of solute adsorption slows down exponentially as the solutes readily occupy the active sites on the adsorbent surface.

$$\frac{dq_t}{dt} = \alpha . exp^{-\beta q_t} \tag{7}$$

q_t is the quantity of pollutant sorption at any time *t* (mg/g)

α is the rate of dye sorption initially (mg/g.min)

β is Elovich constant (g/mg)

When q_t approaches 0, $\frac{dq_t}{dt}$ approaches α

Integrating with limits t(0,t) and q_t (0,q), linearization can be carried out as:

$$q_t = \frac{1}{\beta} \ln[t + t_o] - \frac{1}{\beta} \ln(t_o) \tag{8}$$

Where $t = 1/\alpha\beta$. If $t \gg t_o$, Eq. (8) would be simplified as:

$$q_t = \frac{1}{\beta} \ln[\alpha\beta] + \frac{1}{\beta} lnt \tag{9}$$

The graph of q_t vs ln t helps in determining the heterogeneous nature of the system.[37] α and β can be calculated through the slope and intercept of q_t vs. ln t graph. The greater the values of q_t, α and β, there is more adsorption on the active sites. Thus, greater adsorption capacities mean higher values of these parameters.

In order to check the interaction of dye with the surface of hydrogel, Langmuir and Freundlich isotherms are used.

Langmuir isotherm

Langmuir equation is used to determine the maximum adsorption capacity corresponding to the formation of a monolayer on the adsorbent surface.[38,39]

$$\frac{C_e}{Q_e} = \frac{1}{bQ_{max}} + \frac{C_e}{Q_{max}} \tag{10}$$

where C_e is the equilibrium dye concentration (mg/L);

b is the Langmuir constant (L/mg);

Q_{max} is the maximum monolayer adsorption capacity (mg/g);

Q_e is the equilibrium adsorption capacity (mg/g)

On plotting specific adsorption (C_e/Q_e) against the equilibrium concentration C_e, if a linear plot is obtained, it shows that the adsorption obeys Langmuir isotherm model. The values of Q_{max} and b can be obtained through the slope and intercept, respectively.

The shape of the Langmuir isotherm can be unfavorable, favorable, linear or irreversible depending on the value of a dimensionless separation factor, R_L.

$$R_L = \frac{1}{1 + bC_o} \tag{11}$$

The shape is unfavorable if R_L>1 (in case of desorption); linear if R_L = 1; favorable if R_L lies between 0 and 1 (in case of noticeable adsorption) and irreversible if R_L = 0 (in case of really strong adsorption). Degree of favorability relates to irreversibility of an adsorption system. If R_L approaches 1, it also means the adsorption is reversible in nature.

Freundlich isotherm

$$Q_e = K_f C_e^{1/n} \tag{12}$$

On taking the log on both sides,

$$logQ_e = \log K_f + \frac{1}{n} logC_e \tag{13}$$

where K_f and n are Freundlich constants corresponding to adsorption capacity and adsorption intensity, respectively. Higher values of Q_e, n and K_f indicate higher adsorption capacity of the adsorbent.

Thermodynamic assessment

Enthalpy change (ΔH*), standard entropy change (ΔS*) and Gibbs free energy (ΔG*) are the thermodynamic parameters that should be assessed in order to study the endothermic/exothermic nature and spontaneity of the adsorption process.[40] For the assessment, temperature-dependent adsorption studies need to be carried out. Gibbs free energy change can be calculated using the equation:

$$\Delta G^* = -RTlnK_c \tag{14}$$

where K_c is equilibrium thermodynamic constant which can be calculated by dividing equilibrium adsorption capacity by equilibrium concentration of dye. The value of this parameter depends on temperature. If ΔG* is positive, the value of K_c is less than 1 and is greater than 1 in the other case.

R is ideal gas constant (8.314 J/mol K)

T is absolute temperature (K)

We know that Gibbs free energy is also stated by equation:

$$\Delta G^* = \Delta H^* - T\Delta S^* \tag{15}$$

On equating Eq. (14) and (15),

$$lnK_c = -\frac{\Delta H^*}{RT} + \frac{\Delta S^*}{R} \tag{16}$$

Values of ΔH* and ΔS* are related with K_c through equation (16) and can be obtained by plotting the graph between $\ln K_c$ vs. 1/T and calculating the slope and intercept values.

Results and discussion

Synthesis mechanism

Chitin is a polysaccharide obtained from shells of crustaceans and CS is obtained by its deacetylation using an alkali.[41] Chitin cannot be used directly, as it is immiscible in water at all pH values. On the other hand, CS, which is a glucosamine polymer, can be dissolved in mild acids. For the hydrogel to be synthesized, the other component of interaction apart from CS is MO gum that is composed of different ratios of sugar units like arabinose, galactose, glucuronic acids and rhamnose. The third agent used, i.e. glutaraldehyde, is a saturated dialdehyde that reacts with amines as a crosslinking agent.[42] The synthesis mechanism of CS/MO hydrogel composite is depicted in Figure. 1.

The reaction between chitosan and MO Gum as shown in Figure. 1 occurs due to Maillard linkage between the amino groups of CS and reducing sugar groups of MO Gum. Slight heating leads to loss of water molecules from terminal carbonyl group of MO Gum, leading to $-CH_2NH$ linkage (Maillard linkage). Similar Maillard linkages have been observed between amino groups of glycine, chitosan and sodium caseinate and carbonyl group in glucose, fructose and other such sugar-based complexes.[43,44]

Swelling studies

The effect of nature of the swelling medium on the swelling properties of hydrogel were assessed, and results are presented in Figure 2.

Swelling studies were conducted at three different pH values, i.e., 2.5, 7 and 10.5. All ES% results were obtained in triplicate, and their means were plotted along with the resulting standard deviation values. For comparison of swelling with the synthesized CS/ MO hydrogel, swelling of CS hydrogels and MO powder was also determined. The swelling of CS hydrogels was about 1000–1100% at pH 2.5 and around 900% at pH 7. This happens because CS is a weak polybase that swells more in acidic medium, as there is proton addition of amino groups. At lower pH values, there is repulsion of polymer chains and rupture of intramolecular hydrogen bonding resulting in increased water movement inside the network. MO gum taken individually shows comparatively much less swelling (about 14–20%) in water as compared to chitosan in powdered form. In case of CS/MO hydrogel, at all pH values, swelling attained were almost same. The combination of these two biopolymers, i.e. CS and MO gum, in their hydrogel form resulted in lower swelling of about 200–300%. The available literature on MO gum describes better swelling at neutral pH in distilled water.[45] As CS shows good swelling at acidic and neutral pH and MO gum shows good swelling at neutral pH, the combination shows slightly better swelling at neutral pH (at equilibrium) than acidic and alkaline conditions although the variation is not much. The swelling of the synthesized biopolymeric hydrogels increased with increase in temperature till 40°C. With an increase in temperature, the diffusion of water molecules along with segmental mobility of network increases. This brings about more swelling, as about 200% increase in swelling was observed in the hydrogel on change of temperature from 20°C to 40°C. There is also a screening effect that is caused by decrease of osmotic swelling pressure due to the presence of different ions such as Na^+ and Cl^- ions in solution phase and amino and hydroxyl ions in gel phase.[46,47] Thus, the optimum condition for swelling of the synthesized hydrogel could be classified as temperature within 40°C and swelling medium having minimal salt concentration.

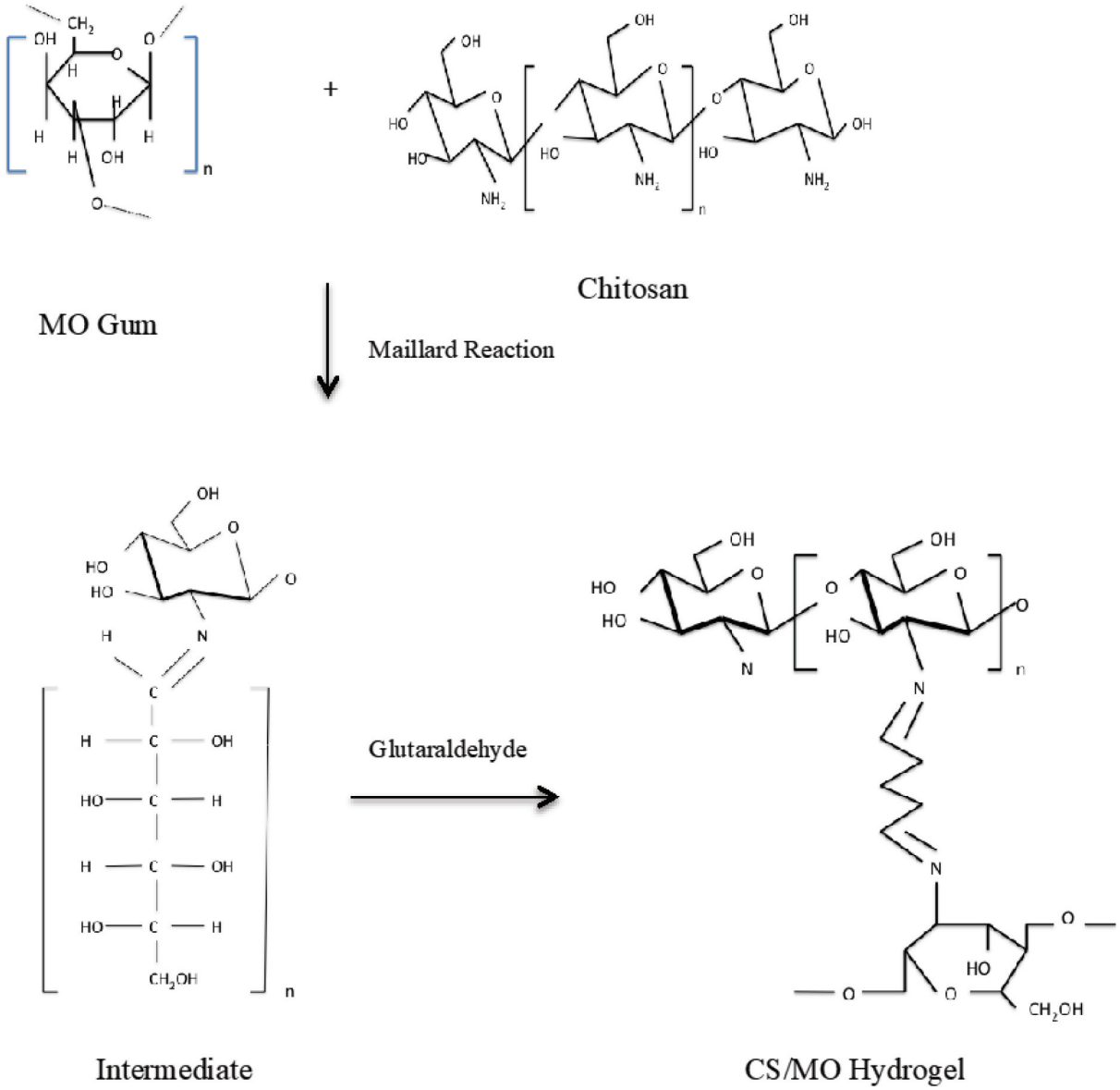


Figure 1. Synthesis mechanism of CS/MO hydrogel.

Characterization of hydrogel

FTIR

The FTIR spectrum of CS/MO hydrogel composite exhibited broad peaks at 3400 cm^{-1} as shown in Figure 3(a), which gives an indication of the presence of O-H stretching due to the presence of hydroxyl group. These groups can also be observed individually in CS hydrogel and MO gum powder. A distinct peak of C-H stretching due to presence of arabinose group was observed at 2924 cm^{-1} for MO gum as well as the formed CS/ MO hydrogel.[48] Carbonyl species peak can be observed at 2164 cm^{-1} in the CS/ MO hydrogel along with C=C alkenes aromatic functional groups at 1604 cm^{-1}. Simultaneously, a peak at 1424 cm^{-1} in the hydrogel confirms the presence of glucouronic acid. Besides this, the peaks near 1070 cm^{-1} in MO gum powder, CS hydrogel and the CS/ MO hydrogel confirm the polysaccharide structure of all three.[49] An interesting observation was the disappearance of the prominent peak at 1650 cm^{-1} in case of CS/MO hydrogel as compared to CS hydrogel. This depicted the linkage of MO gum with NH_2 groups of CS. A new peak appeared at 1604 cm^{-1} showing C-C stretching of benzene ring and ring deformation possibly exhibiting the strong interaction between CS and MO gum.[50]

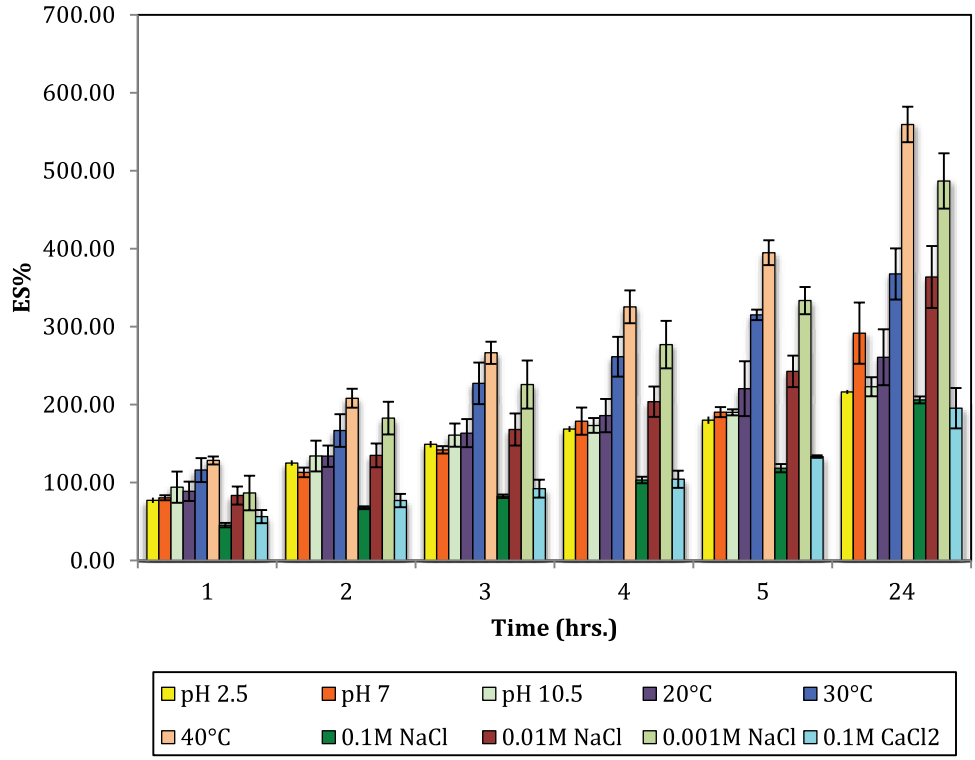


Figure 2. Swelling at different pH, temperature and salt concentrations.

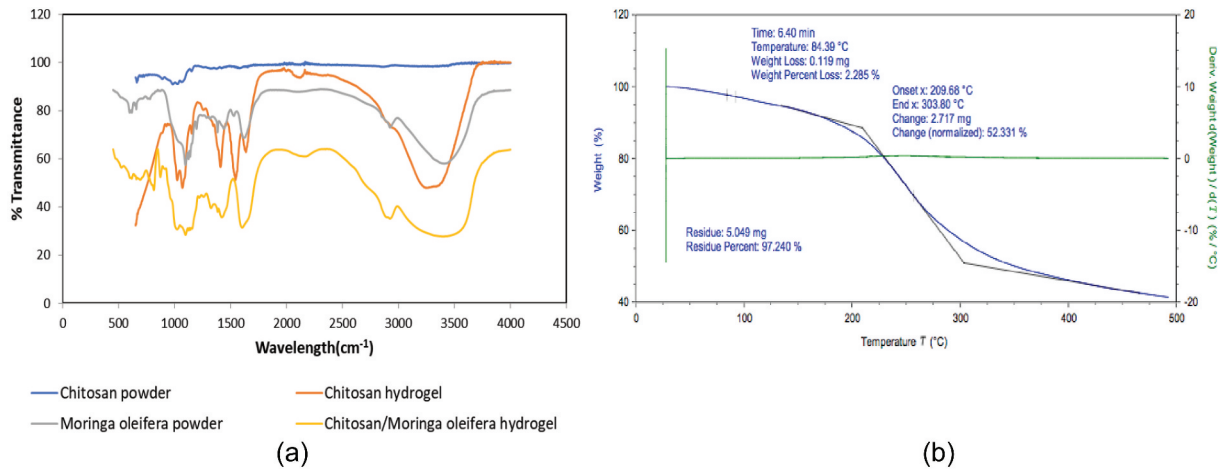


Figure 3. Characterization studies: (a) FTIR and (b) TGA of CS/MO hydrogel.

TGA

A two-stage thermal decomposition occurs for CS/MO hydrogel as depicted in Figure 3(b). A minor first stage, depicting the loss of bound water, occurs at around 84°C. In this stage, 2.28% of weight loss of sample occurs that may be attributed to the moisture present.[50] The major decomposition stage starts at 209°C and complete decomposition occurs at 303°C, leading to a reduction in 52% of quantity initially taken. This could be attributed to the loss of volatile compounds like acetyl compounds, aromatic heterocyclics and CO_2 present in the polymeric network and the conversion of complex molecules such as polysaccharides to acetic acid and other low molecular weight fatty acids.

Dye removal results

Effect of adsorbent dosage on varying dye concentrations

Finding optimum dosages of adsorbent for different dye concentrations is important to understand the removal mechanism and the influence this process has on the amount of dye adsorbed.

The pH was kept as 7 and room temperature (25°C) was used for studying the effect of adsorbent dosage on Congo red dye adsorption as these are the prevalent conditions of usage of the hydrogel based on the gel swelling and considering its utility in the textile industry.[51]

With reference to Figure 4, for obtaining the optimum dosage of adsorbent values, the adsorbent dosages were varied from 0.5 to 4 g/L. It is generally observed that with an increase in adsorbent dosage, an increase in % removal efficiency is observed till a specific value.[52] This could be attributed to increase in the number of active sites available for adsorption. Thereafter, a further increase in dosage of adsorbent does not alter the % removal results. This optimum value after which no improvement is observed in removal efficiency has been termed as the optimum dosage.

After further increase in dosage of adsorbent, the % removal achieved was almost same as shown in Table 2. So, the adsorbent value observed after which the removal remains the same was considered to be the optimum dosage. Thus, as per Figure. 4, it can be clearly observed that for 10, 50, 100 and 1000 mg/L dye concentrations, the optimum dosages are 1, 1.5, 2 and 4 g/L, respectively.

Effect of pH

The pH of the contact solution plays an important role in the adsorption processes protonation/deprotonation of adsorbent occurs at different pH values. The effect of pH was studied by varying the pH of dye solution from 2 to 12 through the graphs of adsorption capacity of hydrogel. The pH of the solutions was adjusted by adding 0.1 M HCl and 0.1 M NaOH. Strong acids and bases can produce a relatively large amount of hydrogen and hydroxide ions and, as a consequence, have marked chemical activity. In addition, very small amounts of strong acids and bases can be used to change the pH of a solution very quickly. On the contrary, buffers are required in large quantities. Thus, in adsorption studies, as usage of buffers was not economically feasible, strong acid and base was used.

Although the optimum dosage worked out to be 4 g/L in case of 1000 mg/L dye concentration, but it had shown rapid removal. Which means that the adsorbent active sites were not fully utilized. So, the 2 g/L adsorbent which was in the form of pellets were cut into smaller fragments and suspended in the dye solution. This process increased the surface area and resulted in

Table 2. % Removal efficiency of dye at different concentrations by varying the adsorbent dosage.

Adsorbent dose	10 mg/L % Removal	50 mg/L % Removal	100 mg/L % Removal	1000 mg/L % Removal
0.5	96.53	78.92	42.49	15.68
1	98.97	93.27	62.64	20.84
1.5	98.6	99.4	82.57	50.38
2		99.37	99.89	82.46
2.5			99.75	83.74
3				86.96
3.5				89.48
4				99.7

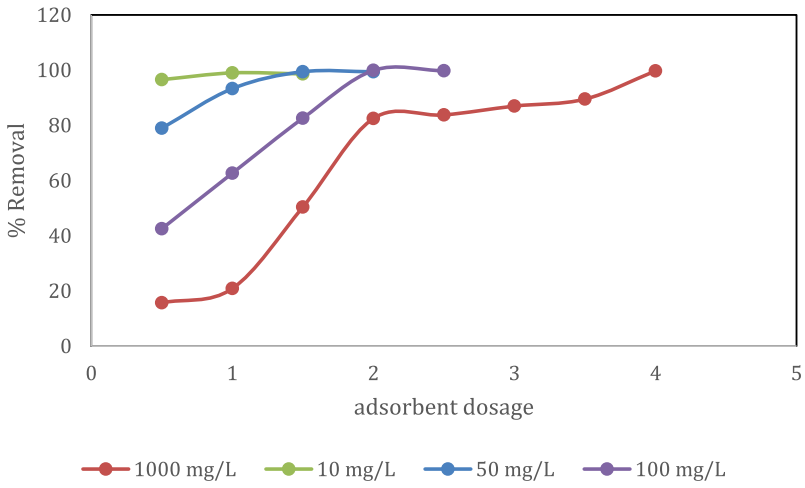


Figure 4. Effect of adsorbent dosage on % removal of dye.

full utility of the active sites of the adsorbent. So, adsorbent dosage of 2 g/L was chosen for conducting the further studies as this worked well at low pH values and all pore spaces could be fully utilized. Moreover, a 4 g/L adsorbent dosage led to some fluctuation in pH of adsorbate due to possibility of some ion exchange mechanism in case of higher dosage of adsorbent. This pH change was not seen in case of 2 g/L of adsorbent.

The experiment was carried out at 25°C temperature, as it is the ambient or room temperature as per International Union of Pure and Applied Chemistry. Because of the presence of chitosan, a decrease in pH leads to an increase in the number of protonated amino groups.[25] However, at higher pH, more OH ions are available that try to repel the anionic Congo red dye molecules, leading to lower adsorption capacity of dye at high pH values that can be seen in Figure. 5. Thus, in case of the present hydrogel, lower pH values are more favorable for Congo red dye removal. Apart from adsorption, some destabilization and settling of dye particles could also be visually observed. These two mechanisms could be attributed to the coagulating nature of MO gum which could further be justified by the study conducted by Ranote et al, 2018, which had illustrated the coagulating and flocculating nature of MO gum.[48] Thus, the combination of CS and MO in the hydrogel had proved to combine the effects of adsorption as well as coagulation, which lead to a high % removal. As CS is protonated in acidic condition, it tends to destabilize the dye particles by charge neutralization.[53] Thus, rapid settling and adsorption of dye particles can be observed at pH 2. At lower pH values ranging from 2 to 6, chitosan can prove to act as an effective adsorbent as well as coagulant. Adsorbent surface may undergo protonation/deprotonation of its surface due to amphoteric nature of water. In order to know whether an adsorbent surface is more acidic or basic, its isoelectric point is determined.[54–56] The isoelectric point or the point of zero charge for CS/MO was experimentally determined to be 5 by method described by Vijayakumar et al, 2012.[53] Thus, the overall surface charge of the hydrogel remains positive at pH<5 because of which there is more electrostatic attraction of anionic dye. The surface charge becomes negative at pH>5 and the adsorption capacity of anionic dye decreases. So, the best state of adsorption in this case for Congo red dye using CS/MO hydrogel is at pH 2.

It is important to note that in case of CS/MO hydrogel, the isoelectric point could be tuned/ changed by slightly varying the composition % of CS and MO gum, as they have varied isoelectric points varying between 7–9 and 1.5–3, respectively. This could prove to be useful if overall surface charge of adsorbent needs to be changed slightly.

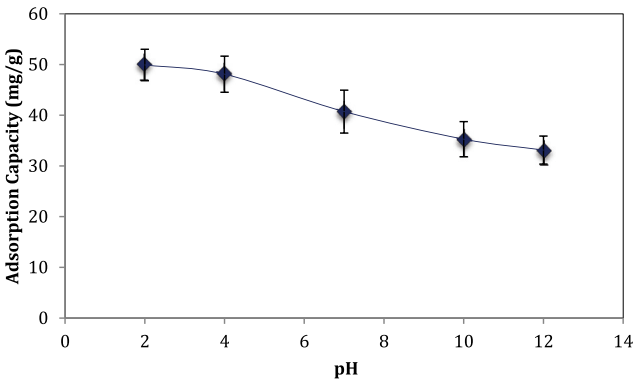


Figure 5. Adsorption capacity of hydrogel for 1000 mg/L dye at different pH till 60 mins.

Effect of contact time and initial dye concentration

With the increase in initial dye concentration from 10 to 1000 mg/L, the equilibrium adsorption capacity increased from 0.49 to 40.72 mg/g (Figure. 6(a)) for adsorption carried out at 25°C temperature, pH = 7

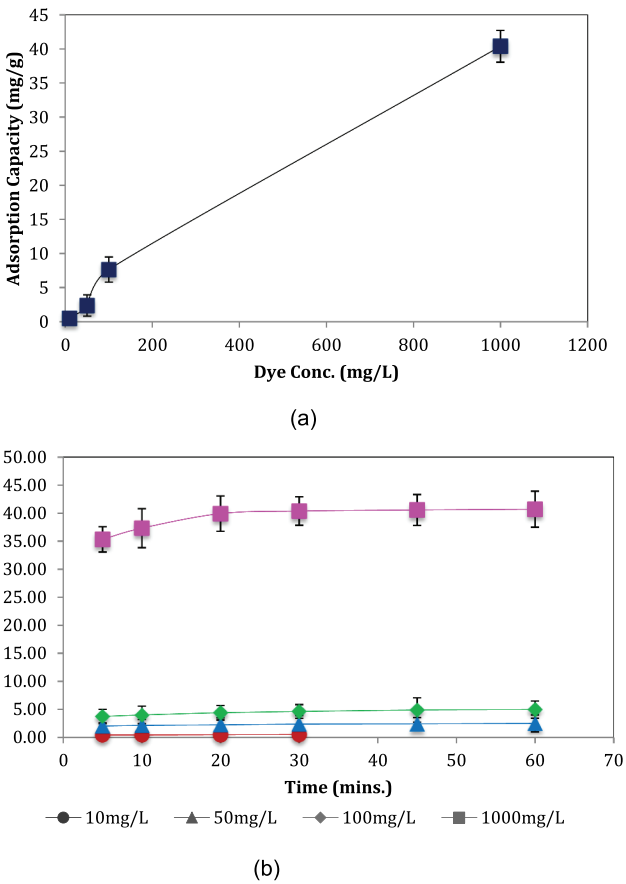


Figure 6. Effect of (a) Initial dye concentration on adsorption capacity and (b) contact time on adsorption capacity.

and adsorbent dosage of 2 g. While studying the effect of pH, it needs a specific mention that Congo red dye is also a well-known pH indicator. At lower pH, i.e. below 5, it changes from its original red color and shifts to dark blue.[51] Though, while conducting the feasibility experiments, the optimum pH was found to be 2, it could not be used for further experiments as the textile industries work with original color of Congo red dye, i.e. red color at a pH between 7 and 10 due to the stability of the color on fabrics. When the pH of the dye is adjusted to be more than 10, again its original color changes from red to wine red. So, deliberately to maintain the standard color of the dye, the pH was set as 7 for conducting further experiments. To support the selection of pH 7, studies relating to adsorption have also been cited in text in the section "**Comparison with other CS based adsorbents**," which have been conducted at pH 7 and room temperature.[2,19,20,57]*

At lower dye concentrations, there are fewer adsorbate particles available for attachment to the adsorbent. As the mass concentration of adsorbate increases, there is more number of collisions of dye molecules with the adsorbent and their moving away from the solution phase.[2] This leads to increase in mass transfer driving force, which is responsible for the increase in the adsorption capacity in case of increase in dye concentration. This shows that the initial dye concentration plays an important role in affecting the adsorption capacity, and it can be concluded that higher the dye concentration, higher the adsorption capacity. The adsorption capacity curves at lower dye concentrations seems to show almost a constant value although the adsorption capacity is increasing. The values are extremely low as compared to values at 1000 mg/L concentration. The % removal of dye decreased in case of high dye concentrations at same adsorbent dosage. This was because adequate pores and surface area were not available for dye adsorption.

The variation in the adsorption capacity of hydrogel with contact time can be observed in Figure. 6(b) for different dye concentrations. It was observed that dye adsorption was rapid till first half an hour of adsorption after which the process slowed down and there was less change in the adsorption capacity values. The % removal of dye occurred quite rapidly with contact. Availability of biosorbent and flocculant leads to this fast removal.

Effect of temperature and thermodynamic evaluation of adsorption

To predict the effect of temperature on adsorption of Congo red dye onto CS/MO hydrogels, thermodynamic studies were carried out at three different temperatures, i.e. 25°, 30° and 40°C. The curve for estimation of thermodynamic parameters was plotted, and equation was used for further calculations. It was observed that when the temperature increased from 25° to 40°C, the adsorption capacity increased from 23.53 mg/g to 24.73 mg/g for 1000 mg/L dye with optimum adsorbent dosage at pH 7 for a contact time of 60 minutes.

The variation in adsorption capacity values is low between temperature varying between 25° and 40°C. The adsorption capacity variation along with the standard deviation values are 23.53 ± 0.16 mg/g to 24.73 ± 0.22 mg/g.

Close values for adsorption capacity varying from 0.37 to about 0.46 mg/g in case of Congo red dye adsorption on varying temperature from 25 to 60°C have also been reported by Litefti et al..[3]

This shows that higher temperature facilitates the adsorption process and equilibrium is achieved in lesser contact time. Further, the thermodynamic parameters were evaluated for determining the spontaneity of the process. The parameters evaluated were free energy change (ΔG), enthalpy change (ΔH) and entropy change (ΔS) using thermodynamic equilibrium constant (K_c) and its variation with temperature. Values obtained for ΔG were −4.92, −4.17 and −3.42 kJ/mol at 298, 303 and 313 K temperatures, respectively, indicating that ΔG increases with increasing incubation temperature. The negative values suggest the feasibility and spontaneity of the reaction, with an increase depicting an increase in spontaneity.[53] The value for ΔS was also negative, indicating a decrease in degree of disorderliness at the solid/liquid interface during the adsorption process. Negative value for ΔH confirms the exothermic nature of the process. Also, the value of ΔH suggests the possibility of physisorption and chemisorption. In case of physisorption, we may expect the dye particles to attach to the gel surface by Vander Waals forces, whereas in case of chemisorption, the bonding is through strong valence forces and desorption becomes difficult. Based on the value of ΔH (−26.88 kJ/mol), the adsorption of Congo red onto the hydrogel proceeded chemically since it falls in the range −20.9–418.4 kJ/mol.[58]

Room temperature (25°C) was used for subsequent experiments as this is the prevalent condition of usage of the hydrogel based on the gel swelling and considering its utility in the textile industry.[51]

Adsorption kinetics

In order to ascertain the fitness of the experimental data with different models, the correlation coefficient or the R^2 values were calculated from graphs of Pseudo first,

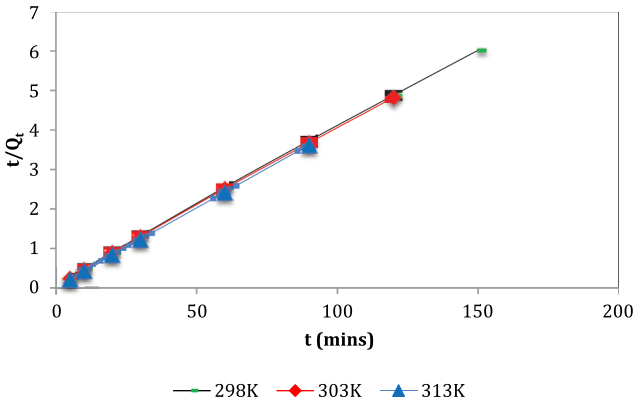


Figure 7. Pseudo-second order kinetics data-interpolation.

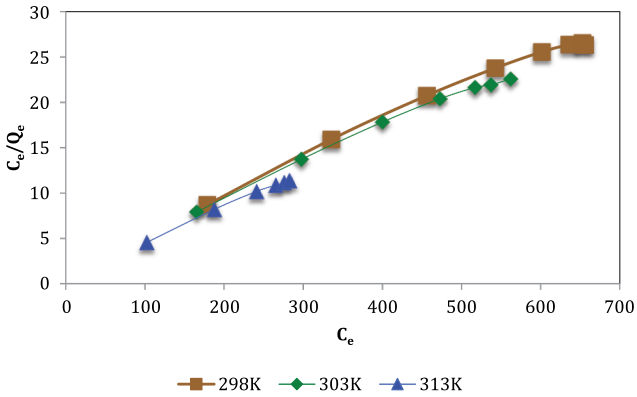


Figure 8. Langmuir Isotherm.

Pseudo second (Figure. 7) and Elovich model, respectively. R^2 close to 1 as shown in Table 3 was obtained for pseudo second-order kinetics as compared to pseudo first-order kinetics and Elovich model.

The R^2 value approaching unity signifies strong physio-chemical interactions between two phases resulting in sorption governed by pseudo-second order kinetics. These are indicative empirical models, and further evaluation needs to be carried out to ascertain if the process is physisorption or chemisorption depending on the thermodynamic parameter evaluation.

Adsorption isotherms

Langmuir and Freundlich isotherms for the adsorption system at different temperatures were plotted for the obtained data. These isotherms are used for interpreting the distribution behavior of adsorbate, i.e. dye on the adsorbent. The R^2 values obtained for Langmuir and Freundlich isotherms (0.99 and 0.84, respectively) concluded that Langmuir isotherm was a better fit model to explain the distribution behavior of the adsorbate (Figure. 8).

On the basis of Langmuir isotherm data plots, the maximum sorption capacity, Q_{max} = 50.25 mg/g, was obtained at pH 2 with temperature 298 K and for dye concentration of 1000 mg/L. This is due to the increasing number of freely protonated amino groups available for adsorption of the dye.[2,18] Since a better fit is obtained for the Langmuir isotherm, it can be interpreted that the hydrogel has a homogeneous surface and dye adsorption occurs as a monolayer. On the contrary, if better fit had been of the Freundlich isotherm, it would have meant that adsorbate distribution was on a heterogeneous surface.

Scanning electron microscopy studies

SEM results as shown in Figure 9 were used to interpret the surface morphology of the hydrogel before and after the dye adsorption. The surface of hydrogel appeared to be rough before adsorption with unevenly distributed pores. The pores were filled with the dye after adsorption, and surface became smoother in appearance. The dark areas on the hydrogel before adsorption process shown by the red circles depict the pore spaces. These areas cannot be seen in the next image. Thus, these SEM images confirm the adsorption process of dye on the hydrogel.

Comparison with other chitosan based adsorbents

Table 4 has been derived after detailed literature review on the adsorption capacities of various chitosan-based hydrogels that have been used for Congo red dye removal. The earlier reported studies have been compared with the hydrogel in this study (row no. 7), so as to mark out its specific relevance and utility.

Table 3. Parameter value evaluation for adsorption of Congo red on CS/MO hydrogel for different isotherms and models.

Temp. (K)	Pseudo first order		Pseudo second order			Elovich model			Langmuir parameters				Freundlich parameters			
	K_1	R^2	Q_{e2}	K_2	R^2	β	Ln(αβ)	R^2	Q_{max}	b	R^2	R_L	n	K_r	R^2	Q_{max}
298	0.021	0.98	25.12	0.39	0.99	0.75	13.73	0.99	27.10	0.0122	0.98	0.075	7.45	9.96	0.84	27.10
303	0.026	0.98	25.12	0.50	0.99	0.78	14.83	0.98	26.95	0.0157	0.99	0.059	7.23	10.13	0.87	26.95
313	0.059	0.98	25.06	1.14	0.99	1.11	23.19	0.95	26.52	0.0425	0.99	0.022	10.11	14.01	0.81	26.52

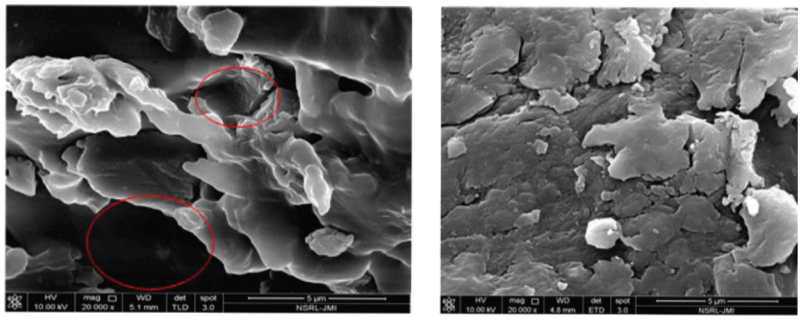


Figure 9. Hydrogel before and after adsorption of Congo red dye.

Table 4. Comparison of various CS based hydrogels used for Congo Red dye removal.

Sl. No.	Adsorbent	pH	Temp. (°C)	% Removal	q_{max}	Order	Isotherm	Ref.
1	Cellulose/CS hydrogel beads	4	25	83.5	40	Pseudo second	Langmuir	[18]
2	Terephthaloyl thiourea crosslinked CS hydrogels	7	25	—	44.25	Pseudo second	Langmuir	[2]
3	Cellulose/CS compound hydrogels	7	25	—	57.44	Pseudo second	Freundlich	[19]
4	CS/Polyacrylamide & graphene oxide composite hydrogels	7	25	94.7	200	Pseudo second	Langmuir	[20]
5	CS/sodium dodecyl sulfate hydrogel beads	5	30	—	209.28	Pseudo second	Langmuir	[16]
6	DNA/ CS hydrogels	7	—	—	12.6	—	—	[21]
7	**CS/ MO hydrogel**	**2**	**25**	**99.82**	**50.25**	**Pseudo second**	**Langmuir**	**This study**

This study used a combination of two biopolymers i.e. CS and MO gum. Moreover, complete % removal of dye was obtained through this hydrogel. Other studies indicated around 84–94% removal signifying superiority of this CS-based hydrogel. Also, though the maximum adsorption could be achieved at pH 2, generally, removal of dye was more than 95% for all pH values. This was subject to condition that the optimum dose of adsorbent was provided for different concentrations of the dye and could be attributed to the sorptive and flocculating nature of the two biopolymers, which are the major components of this hydrogel. The time for removal varied from 20 minutes to 240 minutes depending on concentration of dye and dye solution pH.

Study of interfering agents associated with Congo red to get an idea of real field applicability

Apart from color due to the presence of dye, the other factor of concern in textile industry effluent is high total dissolved solids (TDS) value (Si et al., 2015). Textile industries use common salt and Glauber salt due to which the TDS levels increase. So, in order to investigate the influence of salts in adsorption of Congo red dye, experiments were conducted using 1000 mg/L of dye solution and varying the NaCl concentrations from 0.5 to 1.5% (w/v).

It can be observed from the Figure 10 that the adsorption capacity of hydrogel increased from 40.72 to 48.37 mg/g, whereas the % removal of dye decreased in this case by about 10% when NaCl dosage varied from 0 to 1.5% (w/v) as chloride ions also compete for the attachment on the active sites present in the hydrogel. Although there is no relationship between adsorption capacity and % removal, the reason for the decrease in % removal of dye could be the competition between chloride ions and the dye molecules for binding to the adsorbent surface as both are anionic in nature. Chloride anions existing in the solution may also form complexes with the dye and therefore affect the adsorption process adversely. Thus, chemical precipitation/coagulation of chloride ions before the adsorption of dyes could favor in the complete removal of the dye.

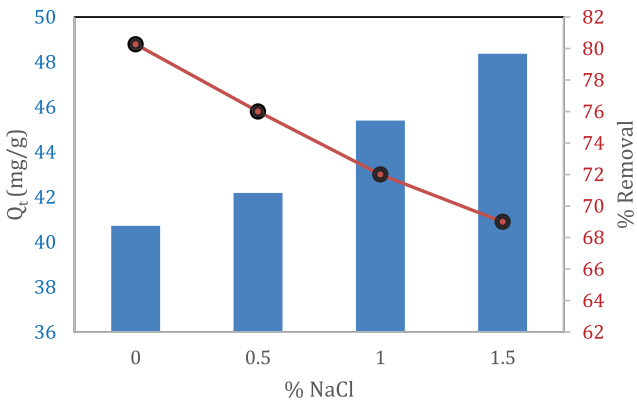


Figure 10. Variation of adsorption capacity and % removal with increasing salt concentrations.

Recovery and reuse

In the times wherein the waste treatment is itself becoming a daunting task, it was proposed to understand the recovery and reuse of this hydrogel to ensure its optimum use. For the aforesaid reason, effort was made to recover the active sites of the adsorbent using HCl or NaOH solutions, the concentration of which was dependent on the nature of the adsorbent.[59] The hydrogels were washed thoroughly with deionized water and then the dye was desorbed using 0.1 M NaOH. Subsequently, the adsorbent samples were washed with slightly acidic water in order to attain protonation of amino/carboxylic groups. The % removal in three consequent cycles were 99%, 90% and 60%. After the third cycle, the pore spaces decreased significantly and the hydrogel could not be further used for sorption purpose.

Conclusion

A novel hydrogel composed of two biopolymers CS and MO gum was synthesized using Maillard reaction under low heating and crosslinking using a very small dosage of 1% glutaraldehyde. The low cost and biocompatible hydrogel showed the presence of amino and hydroxyl groups due to CS, resulting in good hydrophilicity and adsorption ability. As MO gum had a polyuronide structure, the hydrogel combined the advantage of forming a thick stabilizing adsorption layer causing more dye particles to deposit on the surface of hydrogel as compared to CS alone. The chosen anionic azo-dye Congo red dye with concentrations ranging from 10 mg/L to 1000 mg/L could be efficiently adsorbed over a small dosage of hydrogel with a removal efficiency of more than 95% at all tested pH and temperature conditions although with variable contact time. Besides this, these hydrogels are also easy to separate from aqueous streams since they are in the form of pellets unlike the powders usually used for adsorption. These pellets do not dissolve in water as they are crosslinked and settle at the bottom after adsorption. Because of the numerous advantages, these hydrogels show promise for the textile and histological sectors for the removal of anionic dyes.

Acknowledgements

The authors would like to thank Guru Gobind Singh Indraprastha University for the financial assistance provided in the form of Indraprastha Research Fellowship and Faculty Research Grant Scheme, 2020-21 for carrying out the work. The authors would also like to thank Nanotechnology Laboratory, Jamia Millia Islamia University, and SMITA Research Laboratory, Department of Textile Technology, Indian Institute of Technology, Delhi, for providing FTIR, TGA and DSC facilities, respectively, for carrying out the work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Naseem, K.; Farooqi, Z. H.; Begum, R.; Irfan, A. Removal of Congo Red Dye from Aqueous Medium by Its Catalytic Reduction Using Sodium Borohydride in the Presence of Various Inorganic nano-catalysts: A Review. *J. Clean. Prod.* 2018, 187, 296–307. DOI: 10.1016/j.jclepro.2018.03.209.

[2] El-Harby, N. F.; Ibrahim, S. M. A.; Mohamed, N. A. Adsorption of Congo Red Dye onto Antimicrobial Terephthaloyl Thiourea cross-linked Chitosan Hydrogels. *Water Sci. Technol.* 2017, 76(10), 2719–2732. DOI: 10.2166/wst.2017.442.

[3] Litefti, K.; Freire, M. S.; Stitou, M.; González-Álvarez, J. Adsorption of an Anionic Dye (Congo Red) from Aqueous Solutions by Pine Bark. *Sci. Rep.* 2019, 9(1), 16530. DOI: 10.1038/s41598-019-53046-z.

[4] Nawaz, M. S.; Ahsan, M. Comparison of physico-chemical, Advanced Oxidation and Biological Techniques for the Textile Wastewater Treatment. *Alex. Eng. J.* 2014, 53(3), 717–722. DOI: 10.1016/j.aej.2014.06.007.

[5] Pourjavadi, A.; Harzandi, A. M.; Hosseinzadeh, H. Synthesis of a Novel polysaccharide-based Superabsorbent Hydrogel via Graft Copolymerization of Acrylic Acid onto kappa-carrageenan in Air. *Eur. Polym. J.* 2004, 40(7), 1363–1370. DOI: 10.1016/j.eurpolymj.2004.02.016.

[6] Chang, C.; Duan, B.; Cai, J.; Zhang, L. Superabsorbent Hydrogels Based on Cellulose for Smart Swelling and Controllable Delivery. *Eur. Polym. J.* 2010, 46(1), 92–100. DOI: 10.1016/j.eurpolymj.2009.04.033.

[7] Guilherme, M. R.; Aouada, F. A.; Fajardo, A. R.; Martins, A. F.; Paulino, A. T.; Davi, M. F. T.; Rubira, A. F.; Muniz, E. C. Superabsorbent Hydrogels Based on Polysaccharides for Application in Agriculture as Soil Conditioner and Nutrient Carrier: A Review. *Eur. Polym. J.* 2015, 72, 365–385. DOI: 10.1016/j.eurpolmj.2015.04.017.

[8] Cheng, B.; Pei, B.; Wang, Z.; Hu, Q. Advances in chitosan-based Superabsorbent Hydrogels. *RSC Adv.* 2017, 7(67), 42036–42046. DOI: 10.1039/c7ra07104c.

[9] Cui, H. F.; Wu, W. W.; Li, M. M.; Song, X.; Lv, Y.; Zhang, -T.-T. A Highly Stable Acetylcholinesterase Biosensor Based on chitosan-TiO2-graphene Nanocomposites for Detection of Organophosphate Pesticides. *Biosens. Bioelectron.* 2018, 99, 223–229. DOI: 10.1016/j.bios.2017.07.068.

[10] Babu, V. R.; Hosamani, K. M.; Aminabhavi, T. M. Preparation and in-vitro Release of Chlorothiazide Novel pH-sensitive chitosan-N,N′-dimethylacrylamide

semi-interpenetrating Network Microspheres. *Carbohydr. Polym.* 2008, 71(2), 208–217. DOI: 10.1016/j.carbpol.2007.05.039.

[11] Bashir, S.; Teo, Y. Y.; Ramesh, S.; Ramesh, K. Synthesis, Characterization, Properties of N-succinyl chitosan-g-poly (Methacrylic Acid) Hydrogels and in Vitro Release of Theophylline. *Polym.* 2016, 92, 36–49. DOI: 10.1016/j.polymer.2016.03.045.

[12] Jana, S.; Sen, K. K. Chitosan — Locust Bean Gum Interpenetrating Polymeric Network Nanocomposites for Delivery of Aceclofenac. *Int. J. Biol. Macromol.* 2017, 102, 878–884. DOI: 10.1016/j.ijbiomac.2017.04.097.

[13] Cui, L.; Jia, J.; Guo, Y.; Liu, Y.; Zhu, P. Preparation and Characterization of IPN Hydrogels Composed of Chitosan and Gelatin cross-linked by Genipin. *Carbohydr. Polym.* 2014, 99, 31–38. DOI: 10.1016/j.carbpol.2013.08.048.

[14] Parwani, L.; Bhatnagar, M.; Bhatnagar, A.; Sharma, V.; Sharma, V. Evaluation of Moringa Oleifera Seed biopolymer-PVA Composite Hydrogel in Wound Healing Dressing. *Iran. Polym. J. (Eng Ed).* 2016, 25(11), 919–931. DOI: 10.1007/s13726-016-0479-8.

[15] Panda, D.; Choudhury, N. S. K.; Yedukondalu, M.; Si, S.; Gupta, R. Evaluation of Gum of *Moringa Oleifera* as a Binder and Release Retardant in Tablet Formulation. *Ind. J. Pharm. Sc.* 2008, 70(5), 614. DOI: 10.4103/0250-474X.45400.

[16] Vigneshwaran, S.; Karthikeyan, P.; Sirajudheen, P.; Meenakshi, S. Optimization of Sustainable chitosan/Moringa. Oleifera as Coagulant Aid for the Treatment of Synthetic Turbid water–A Systemic Study. *Environ. Chem. Ecotox.* 2020, 2, 132–140. DOI: 10.1016/j.enceco.2020.08.002.

[17] Naqvi, S. M. K.; Khan, Z.; Mirza, E. H.; Chandio, A.; Manzoor, F.; Niaz, R.; Khan, A. A.; AlKhureif, A. A. Fabrication and Characterization of Polyvinyl alcohol/chitosan/moringa-extract Hydrogel Patch for wound-healing Applications. *Mat. Exp.* 2021, 11(1), 107–115. DOI: 10.1166/mex.2021.1895.

[18] León, O.; Muñoz-Bonilla, A.; Soto, D.; Pérez, D.; Rangel, M.; Colina, M.; Fernández-García, M. Removal of Anionic and Cationic Dyes with Bioadsorbent Oxidized Chitosans. *Carbohydr. Polym.* 2018, 194, 375–383. DOI: 10.1016/j.carbpol.2018.04.072.

[19] Tu, H.; Yu, Y.; Chen, J.; Shi, X.; Zhou, J.; Deng, H.; Du, Y. Highly cost-effective and high-strength Hydrogels as Dye Adsorbents from Natural Polymers: Chitosan and Cellulose. *Polym. Chem.* 2017, 8(19), 2913–2921. DOI: 10.1039/C7PY00223H.

[20] Sami, A. J.; Butt, Y. N.; Nasar, S. Elimination of a Carcinogenic Anionic Dye Congo Red from Water using Hydrogels Based on Chitosan, Acrylamide and Graphene Oxide. *J. Bioprocess Biotech.* 2018. DOI: 10.4172/2155-9821.1000334.

[21] Chan, K.; Morikawa, K.; Shibata, N.; Zinchenko, A. Adsorptive Removal of Heavy Metal Ions, Organic Dyes, and Pharmaceuticals by DNA–chitosan Hydrogels. *Gels.* 2021, 7(3), 112. DOI: 10.3390/gels7030112.

[22] Ramakrishnan, R. K.; Padil, V. V. T.; Wacławek, S.; Černík, M.; Varma, R. S. Eco-friendly and Economic, Adsorptive Removal of Cationic and Anionic Dyes by bio-based Karaya gum—chitosan Sponge. *Polym.* 2021, 13(2), 251. DOI: 10.3390/polym13020251.

[23] Ranote, S.; Kumar, D.; Kumari, S.; Kumar, R.; Chauhan, G. S.; Joshi, V. Green Synthesis of Moringa Oleifera gum-based Bifunctional Polyurethane Foam Braced with Ash for Rapid and Efficient Dye Removal. *Chem. Engg. J.* 2019, 361, 1586–1596. DOI: 10.1016/j.cej.2018.10.194.

[24] Ranote, S.; Chauhan, G. S.; Joshi, V. Etherified Moringa Oleifera Gum as Rapid and Effective Dye Adsorbents. *Chem. Engg. J.* 2020, 387, 124055. DOI: 10.1016/j.cej.2020.124055.

[25] Ravikumar, K.; Udayakumar, J. Moringa Oleifera Gum Composite a Novel Material for Heavy Metals Removal. *Int. J. Environ. Anal. Chem.* 2019, 1–41. DOI: 10.1080/03067319.2019.1686142.

[26] Ahmad, S.; Manzoor, K.; Purwar, R.; Ikram, S. Morphological and Swelling Potential Evaluation of Moringa Oleifera gum/poly (Vinyl Alcohol) Hydrogels as a Superabsorbent. *ACS omega.* 2020, 5(29), 17955–17961. DOI: 10.1021/acsomega.0c01023.

[27] Singh, B.; Kumar, A. Network Formation of Moringa Oleifera Gum by Radiation Induced Crosslinking: Evaluation of Drug Delivery, Network Parameters and Biomedical Properties. *Int. J. Bio. Macromol.* 2018, 108, 477–488. DOI: 10.1016/j.ijbiomac.2017.12.04.

[28] Teimouri, A.; Nasab, S. G.; Vahdatpoor, N.; Habibollahi, S.; Salavati, H.; Chermahini, A. N. Chitosan /Zeolite Y/Nano ZrO2 Nanocomposite as an Adsorbent for the Removal of Nitrate from the Aqueous Solution. *Int. J. Biol. Macromol.* 2016, 93, 254–266. DOI: 10.1016/j.ijbiomac.2016.05.089.

[29] Ganji, F.; Vasheghani-Farahani, S.; Vasheghani-Farahani, E. Theoretical Description of Hydrogel Swelling: A Review. *Iran. Polym. J.* 2010, 19(5), 375–398.

[30] Sievers, J.; Sperlich, K.; Stahnke, T.; Kreiner, C.; Eickner, T.; Martin, H.; Guthoff, R. F.; Schünemann, M.; Bohn, S.; Stachs, O. Determination of Hydrogel Swelling Factors by Two Established and a Novel non-contact Continuous Method. *J. Appl. Polym. Sci.* 2021, 138(18), 50326. DOI: 10.1002/app.50326.

[31] Huang, Z. G.; Wang, T.; Yi, H. Y.; Li, X. Study on the Adsorption of Methylene Blue from Dye Wastewater by Humulus Japonicus Leaves. In: *E3S Web of Conferences*, Harbin, China, 25-27 December 2020; 2021. DOI: 10.1051/e3sconf/202123603028.

[32] Moussout, H.; Ahlafi, H.; Aazza, M.; Maghat, H. Critical of Linear and Nonlinear Equations of pseudo-first Order and pseudo-second Order Kinetic Models. *Karbala Int. J. Modern Sci.* 2018, 4(2), 244–254. DOI: 10.1016/j.kijoms.2018.04.001.

[33] Ho, Y. S. Review of second-order Models for Adsorption Systems. *J. Hazard. Mater.* 2006, 136(3), 681–689. DOI: 10.1016/j.jhazmat.2005.12.043.

[34] Sağ, Y.; Aktay, Y. Kinetic Studies on Sorption of Cr(VI) and Cu(II) Ions by Chitin, Chitosan and Rhizopus Arrhizus. *Biochem. Eng. J.* 2002, 12(2), 143–153. DOI: 10.1016/S1369-703X(02)00068-2.

[35] Wu, F. C.; Tseng, R. L.; Juang, R. S. Characteristics of Elovich Equation Used for the Analysis of Adsorption Kinetics in dye-chitosan Systems. *Chem. Eng. J.* 2009, 150(2–3), 366–373. DOI: 10.1016/j.cej.2009.01.014.

[36] Kajjumba, G. W.; Aydın, S.; Güneysu, S. Adsorption Isotherms and Kinetics of Vanadium by Shale and Coal Waste. *Adsorp. Sci. Technol.* 2018, 36(3–4), 936–952. DOI: 10.1177/0263617417733586.

[37] Ho, Y. S.; McKay, G. Application of Kinetic Models to the Sorption of Copper (II) on to Peat. *Adsorp. Sci. Technol.* 2002, 20(8), 797–815. DOI: 10.1260/026361702321104282.

[38] Saruchi, K. V.; Kumar, V. Adsorption Kinetics and Isotherms for the Removal of Rhodamine B Dye and Pb +2 Ions from Aqueous Solutions by a Hybrid ion-exchanger. *Arab. J. Chem.* 2019, 12(3), 316–329. DOI: 10.1016/j.arabjc.2016.11.009.

[39] Mahmoud, M. A. Kinetics and Thermodynamics of Aluminum Oxide Nanopowder as Adsorbent for Fe (III) from Aqueous Solution. *Beni-Suef Univ. J. Appl. Sci.* 2015, 4(2), 142–149. DOI: 10.1016/j.bjbas.2015.05.008.

[40] Zhao, Y.; Zhao, X.; Deng, J.; He, C. Utilization of chitosan–clinoptilolite Composite for the Removal of Radio cobalt from Aqueous Solution: Kinetics and Thermodynamics. *J. Radioanal. Nucl. Chem.* 2016, 308 (2), 701–709. DOI: 10.1007/s10967-015-4475-9.

[41] Islam, S.; Bhuiyan, M. A. R.; Islam, M. N. Chitin and Chitosan: Structure, Properties and Applications in Biomedical Engineering. *J. Polym. Environ.* 2017, 25 (3), 854–866. DOI: 10.1007/s10924-016-0865-5.

[42] Sehmi, S. K.; Allan, E.; MacRobert, A. J.; Parkin, I. The Bactericidal Activity of glutaraldehyde-impregnated Polyurethane. *Microbiol. Open.* 2016, 55, 891–897. DOI:10.1002/mbo3.378.

[43] Gu, F.; Kim, J. M.; Hayat, K.; Xia, S.; Feng, B.; Zhang, X. Characteristics and Antioxidant Activity of Ultrafiltrated Maillard Reaction Products from a casein-glucose Model System. *Food Chem.* 2009, 117 (1), 48–54. DOI: 10.1016/j.foodchem.2009.03.074.

[44] Lund, M. N.; Ray, C. A. Control of Maillard Reactions in Foods: Strategies and Chemical Mechanisms. *J. Agric. Food Chem.* 2017, 65(23), 4537–4552. DOI: 10.1021/acs.jafc.7b00882.

[45] Jarald, E. E.; Sumati, S.; Edwin, S.; Ahmad, S.; Patni, S.; Daud, A. Characterization of Moringa Oleifera Lam. Gum to Establish It as a Pharmaceutical Excipient. *Indian J. Pharma. Edu. Res.* 2012, 46(3), 211–216.

[46] Gupta, V. N.; Shivakumar, H. G. Investigation of Swelling Behavior and Mechanical Properties of a pH-Sensitive Superporous Hydrogel Composite. *Iran J. Pharm. Res.* 2012, 11(2), 481–493. DOI: 10.22037/ijpr.2012.1097.

[47] Khushbu,; Warkar, S. G.; Kumar, A. Synthesis and Assessment of Carboxymethyl Tamarind Kernel Gum Based Novel Superabsorbent Hydrogels for Agricultural Applications. *Polym. J.* 2019, 2019, 182. DOI: 10.1016/j.polymer.2019.121823.

[48] Ranote, S.; Ram, B.; Kumar, D.; Chauhan, G. S.; Joshi, V. Functionalization of Moringa Oleifera Gum for Use as Hg2+ Ions Adsorbent. *J. Environ. Chem. Eng.* 2018, 6 (2), 1805–1813. DOI: 10.1016/j.jece.2018.02.032.

[49] Gipson, K.; Stevens, K.; Brown, P.; Ballato, J. Infrared Spectroscopic Characterization of Photoluminescent Polymer Nanocomposites. *J. Spectrosc.* 2015, 2015, 1–9. DOI: 10.1155/2015/489162.

[50] Pooresmaeil, M.; Namazi, H. Application of polysaccharide-based Hydrogels for Water Treatments. *Hydro. Based on Nat. Poly.* 2019, 411–455. DOI: 10.1016/B978-0-12-816421-1.00014-8.

[51] Taher, T.; Rohendi, D.; Mohadi, R.; Lesbani, A. Congo Red Dye Removal from Aqueous Solution by acid-activated Bentonite from Sarolangun: Kinetic, Equilibrium, and Thermodynamic Studies. *Arab. J. Basic Appl. Sci.* 2019, 26 (1), 125–136. DOI: 10.1080/25765299.2019.1576274.

[52] Sumalatha, S.; Rao, P. P. STUDIES ON BIO-ADSORPTION OF Cd (II) FROM AQUEOUS SOLUTIONS. *Editorial Board.* 2015, 4(11), 116.

[53] Patel, H.; Vashi, R. T. Removal of Congo Red Dye from Its Aqueous Solution Using Natural Coagulants. *J. Saudi Chem. Soc.* 2012, 16(2), 131–136. DOI: 10.1016/j.jscs.2010.12.003.

[54] Wechsler, M. E.; Dang, H. K. H. J.; Dahlhauser, S. D.; Simmonds, S. P.; Reuther, J. F.; Wyse, J. M.; VandeWalle, A. N.; Anslyn, E. V.; Peppas, N. A. Nanogel Receptors for High Isoelectric Point Protein Detection: Influence of Electrostatic and Covalent polymer-protein Interactions. *Chem. Commun.* 2020, (2020)(56), 6141–6144. DOI: 10.1039/d0cc02200d.

[55] Wiśniewska, M.; Sternik, D.; Nowicki, P. Adsorption, Viscosity and Thermal Behaviour of Nanosized Proteins with Different Internal Stability Immobilised on the Surface of Mesoporous Activated Biocarbon Obtained from the Horsetail Herb Precursor. *Appl. Nanosci. (Switzerland).* 2021. DOI: 10.1007/s13204-021-01759-x.

[56] Vijayakumar, G.; Tamilarasan, R.; Dharmendirakumar, M. Adsorption, Kinetic, Equilibrium and Thermodynamic Studies on the Removal of Basic Dye Rhodamine-B from Aqueous Solution by the Use of Natural Adsorbent Perlite. *J. Mater. Environ. Sci.* 2012, 3(1), 157–170.

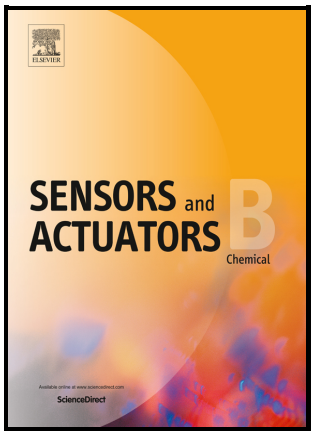
[57] Sari, A.; Tuzen, M. Biosorption of Total Chromium from Aqueous Solution by Red Algae (Ceramium Virgatum): Equilibrium, Kinetic and Thermodynamic Studies. *J. Hazard. Mater.* 2008, 160(2–3), 349–355. DOI: 10.1016/j.jhazmat.2008.03.005.

[58] Li, M.; Wang, Z.; Li, B. Adsorption Behaviour of Congo Red by cellulose/chitosan Hydrogel Beads Regenerated from Ionic Liquid. *Desalin. Water Treat.* 2016, 57, 16970–16980. DOI: 10.1080/19443994.2015.1082945.

[59] Chatterjee, S.; Chatterjee, T.; Lim, S. R.; Woo, S. H. Effect of the Addition Mode of Carbon Nanotubes for the Production of Chitosan Hydrogel core–shell Beads on Adsorption of Congo Red from Aqueous Solution. *Bioresour. Technol.* 2011, 102(6), 4402–4409. DOI: 10.1016/j.biortech.2010.12.117.

Journal Pre-proof

Tyrosinase-functionalized gold nanoparticle-tailored ultrasensitive nanosensing probe for hazardous and nutritional phenolic compounds



Jasmeen Dhiman, Kalyan Vaid, Treesa Johns, Ruchika Maurya, Mahima Arora, Ankita Negi, Ritika Gupta, Mrinmoy Misra, Ki-Hyun Kim, Vanish Kumar

PII: S0925-4005(22)01076-0

DOI: https://doi.org/10.1016/j.snb.2022.132434

Reference: SNB132434

To appear in: *Sensors and Actuators: B. Chemical*

Received date: 24 March 2022
Revised date: 8 July 2022
Accepted date: 25 July 2022

Please cite this article as: Jasmeen Dhiman, Kalyan Vaid, Treesa Johns, Ruchika Maurya, Mahima Arora, Ankita Negi, Ritika Gupta, Mrinmoy Misra, Ki-Hyun Kim and Vanish Kumar, Tyrosinase-functionalized gold nanoparticle-tailored ultrasensitive nanosensing probe for hazardous and nutritional phenolic compounds, *Sensors and Actuators: B. Chemical,* (2022) doi:https://doi.org/10.1016/j.snb.2022.132434

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier.

Tyrosinase-functionalized gold nanoparticle-tailored ultrasensitive nanosensing probe for hazardous and nutritional phenolic compounds

Jasmeen Dhiman[a], Kalyan Vaid[a,b,c], Treesa Johns[a], Ruchika Maurya[a], Mahima Arora[a], Ankita Negi[a], Ritika Gupta[a], Mrinmoy Misra[d], Ki-Hyun Kim[e,*], Vanish Kumar[a,f*]

[a]National Agri-Food Biotechnology Institute (NABI), S.A.S. Nagar, Punjab, 140306, India

[b]Centre for Nanoscience and Nanotechnology, Panjab University, Chandigarh,160014, India

[c]CSIR-Central Scientific Instruments Organisation, Sector 30-C, Chandigarh, 160030, India

[d]Mechatronics Engineering Department, School of Automobile, Mechanical, and Mechatronics, Manipal University, Jaipur, India

[e]Department of Civil and Environmental Engineering, Hanyang University, 222 Wangsimni-ro, Seoul, 04763, Korea

[f]Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala 147004

Journal Pre-proof

Abstract

Herein, a sensitive, selective, and facile sensing approach was developed by tailoring gold nanoparticles (AuNPs) to detect phenolic compounds in water and food/plant products across a range of concentrations. Four different types of phenolic compounds selective sensing probes were developed by varying the capping amounts (µg/mL) of tyrosinase (Ty) on the AuNP surface such as 5, 10, 50, and 100 µg/mL. Accordingly, an ultralow limit of detection (LOD) for phenolic compounds was achieved at 0.01 ppb (along with the upper bound detection of 50,000 ppm) via spectroscopic methodology. Interestingly, the absorbance of the developed sensing probe increased in the presence of phenolic compounds due to separation of Cu (present in Ty) from the proximity of AuNPs. A novel

sensing mechanism has been proposed in light of interaction between enzyme substrate and copper of Ty with gold nanoparticles. Moreover, the applicability of the developed sensing probe is successfully validated using the tea leaves samples.

Keywords: Gold nanoparticles; sensing; polyphenols; biosensor; enzyme; total phenolic content.

*****Correspondence: vanish.saini01@gmail.com; (V. Kumar); kkim61@hanyang.ac.kr (K.-H. Kim)

TOC

1. **Introduction**
2. **MATERIALS AND METHODS**
    - **2.1 Materials**
    - **2.2 Synthesis and characterization of the Ty-AuNP probe**
    - **2.3 Colorimetric detection of phenolic compounds by Ty-AuNPs**
    - **2.4 Polyphenol extraction from tea leaves**
3. **RESULTS AND DISCUSSION**
    - **3.1 Characterization of AuNPs**
    - **3.2 Sensing performance of the Ty-AuNP probe**
    - **3.3 Sensing mechanism of the Ty-AuNP probe**
    - **3.4 Detection of polyphenols in real samples**
    - **3.5 Specificity study of the Ty-AuNP probe**
    - **3.6 Reusability study**
4. **Conclusion**
5. **References**

Journal Pre-proof

1. INTRODUCTION

Phenolic compounds are one of the most studied classes of organic compounds and contain at least one phenol moiety in their chemical structure. These compounds exist in a wide range of chemical structures, from simple (e.g., gallic acid) to complex molecules (e.g., proanthocyanidins) [1]. On the basis of origin, they can also be classified as natural or synthesized/industrial phenolic compounds [2]. Generally, natural phenols (especially polyphenols) are beneficial for human health. Some of them are also categorized as nutraceuticals, as they can help improve the nutritional quality of food [1, 3]. Polyphenols can exhibit anti-oxidant, anti-inflammatory, anti-allergic, and anti-microbial properties [1, 3, 4].

Daily intake of flavonoids (1500 mg/day) and proanthocyanidins (≥229 mg/day) can reduce the risk of bone fracture and renal failure, respectively, in humans [5, 6]. Several types of fruits, vegetables, and beverages (e.g., wines) serve as sources of high total phenolic content (TPC). For example, the estimated TPC in green pepper, yellow tea, and mulberries is reported to be 3050, 6600, and 15200 ppm, respectively [7-9]. Interestingly, TPC is a prime quality assurance parameter for various food items, such as tea and wine [1, 10, 11]. On the other hand, some phenolic compounds (e.g., chlorophenols, nitrophenols, aminophenols, and chlorocatechols) are listed as pollutants by the European Union (EU) and the United States Environmental Protection Agency (USEPA), with a permissible limit of 0.5 ppm [12-14]. These phenolic compounds can cause acute and chronic diseases and infections in humans and animals [12-15].

For these reasons, quantification of TPC is essential for grading of food and plant products and environmental monitoring [14]. The Folin-Ciocalteu (FC) test is one of the most commonly known methods used for such determination [16]. However, it suffers from several shortcomings, such as lengthy procedural steps, complex methodology, limited sensitivity, and reactivity against vitamins,

Journal Pre-proof

proteins, and thiols [17, 18]. Likewise, as other analytical methods, such as high-performance liquid chromatography (HPLC) and gas chromatography-mass spectroscopy (GC-MS), also suffer from disadvantages (e.g., expensive set-ups, complex operations, and tedious procedures), their use as a point-of-care detection system is limited to a certain degree [14]. Hence, the potential of enzymatic biosensors is recognized through multiple advantages, such as low manufacturing costs and facile fabrication, for simple, specific, rapid, and economical monitoring of TPC [19].

Tyrosinase (Ty), also known as phenoloxidase, phenolase, or catechol oxidase, is one of the most well-known members of the polyphenol oxidase (PPO) family and can be used for fabrication of enzymatic biosensors to detect TPC [19, 20]. Ty is a multi-functional copper (Cu)-containing metalloenzyme that catalyzes oxidation of phenolic compounds to their respective quinones [19, 21-23]. Immobilization of Ty on a transducer (e.g., metal nanoparticles, carbon materials, and polymers) surface has been achieved by diverse methodologies, e.g., physical adsorption and covalent cross-linking [19, 21, 24, 25]. Gold nanoparticles (AuNPs) are an excellent choice for fabrication of enzymatic biosensors due to facile enzyme immobilization and surface plasmon resonance (SPR) characteristics [26]. In general, the AuNPs exhibit red-wine color due to SPR effect, although their color can change to dark blue/purple with any physical/chemical instabilities [27]. Thus, colorimetric sensors can be fabricated by employing the SPR effects of AuNPs. It should be noted that the color and light absorption of AuNPs are highly dependent on its size such that the SPR peak shifted towards higher wavelength with increase in the size of AuNPs [28]. Accordingly, the red color is more prominent in small sized AuNPs, while blue color is more common for bigger AuNPs. Till now, the SPR of AuNPs has been explored for the sensing of diverse analytes, e.g., metabolites, drugs, pesticides, and metal contaminants [29-31]. Moreover, the AuNPs are also found to be effective for the bioimaging and drug delivery applications [28].

Journal Pre-proof

Herein, we designed a tailor-made Ty-capped AuNP (Ty-AuNP) probe to sense varying ranges of catechol. The amount of Ty was controlled by varying the ratio of enzyme to AuNPs during synthesis of Ty-AuNP biocomposites to adjust the sensing ranges and sensitivities of the developed probes. The developed technique should be beneficial for determination of phenolic compounds in various food and environmental samples, irrespective of their diversity (e.g., type and content). Moreover, unlike conventional methods, the designed probe is specific for detection of phenols due to the chosen enzyme [22]. Specifically, four types of Ty-AuNP sensing probes were designed and standardized by varying the Ty concentration (5, 10, 50, or 100 µg/mL) during AuNP-synthesis. These probes were named Ty-AuNPs (5), Ty-AuNPs (10), Ty-AuNPs (50), and Ty-AuNPs (100), respectively. The probes were used for colorimetric detection of phenolic compounds from 0.01 ppb to 50,000 ppm with high selectivity. In addition, a novel sensing mechanism has been discussed based on the interactions between the Cu present in the Ty enzyme and AuNPs. To the best of our knowledge, this is the first report on a custom-made probe to sense TPC based on the interactions between Ty and AuNPs.

2. MATERIALS AND METHODS

2.1 Materials

Chloroauric acid ($HAuCl_4$) was obtained from Central Drug House (CDH) (P) Ltd, India. Trisodium citrate, sodium chloride, sodium phosphate monobasic monohydrate, sodium phosphate dibasic, pyrocatechol, Folin-Ciocalteu reagent, sodium carbonate, aluminium chloride, cetrimonium bromide (CTAB) glutathione (GSH), and mushroom Ty were procured from Sigma-Aldrich, USA. The Pierce BCA Protein Assay Kit and hexane were purchased from Thermo Fisher Scientific, USA and Finar Limited, India, respectively. All chemicals used were of analytical grade, and the aqueous solutions were prepared in Milli-Q (MQ) water with 18 MΩ.cm resistivity.

Journal Pre-proof

2.2 Synthesis and characterization of the Ty-AuNP probe

Citrate-AuNPs were prepared using the Turkevich reduction method, where trisodium citrate acts as a reducing and stabilizing agent for gold salt ($HAuCl_4$) and AuNPs, respectively [32]. The size of AuNPs was controlled by varying the concentration of reducing agent [33]. Out of synthesized AuNPs, the NPs with diameter ranges of 9, 20, 28, 36, and 45 nm were used to develop the sensing probes. The size of AuNPs was examined through dynamic light scattering (DLS). The sensing probe was produced by immobilizing Ty onto synthesized AuNPs via physical adsorption. For this, aqueous AuNPs were added dropwise into Ty solution at a 9:1 (v/v) ratio with continuous overnight stirring at 4 °C. The synthesized Ty-AuNPs were washed by 60 min of centrifugation at 8000 rcf. Out of the synthesized AuNPs, 9 nm sized AuNPs were found to be optimum for the generation of sensing signals against catechol. Accordingly, 9 nm sized AuNPs were used for the synthesis of four Ty-AuNPs probes, i.e., Ty-AuNPs (5), Ty-AuNPs (10), Ty-AuNPs (50), and Ty-AuNPs (100) by adding 5, 10, 50, and 100 µg/mL of Ty, respectively. The citrate-capped silver nanoparticles (AgNPs) (synthesized by chemical reduction method) were also explored for the possibilities to develop AgNPs-based sensing probe against catechol. The Ty solution used in current studies was prepared in PBS buffer with pH 7.0. The synthesized Ty-AuNPs were characterized by ultraviolet-visible (UV-vis) spectroscopy (Shimadzu UV-2700), dynamic light scattering; DLS (Malvern Zetasizer Nano-ZS), energy-dispersive X-ray analysis; EDX (JSM 7610FPlus), and field emission scanning electron microscopy; FESEM (JSM 7610FPlus). The quantification of immobilized Ty on AuNPs was performed by BCA protein assay.

2.3 Spectroscopic and colorimetric detection of phenolic compounds by Ty-AuNPs

Journal Pre-proof

For detection of catechol, Ty-AuNPs were mixed with different concentrations of catechol solutions and incubated for 5 min (first step). Then, an increase in the absorbance of Ty-AuNPs was observed (in comparison to control Ty-AuNPs), which allowed quantitation of catechol. The absorbance of Ty-AuNPs increased gradually with increase in the levels of catechol. Accordingly, a standard curve (concentration of catechol vs absorbance) was drawn to quantify the levels of catechol in unknown samples. Moreover, the level of catechol was further verified by measuring localized SPR (LSPR) peak shifts (or by the naked eye) after addition of NaCl (0.5 M NaCl was added after the first step, and the solution was incubated for another 5 min). A 1:5 (v/v) ratio solution of Ty-AuNPs to catechol was used for catechol detection as it showed the best associated AuNP SPR peak.

2.4 Polyphenol extraction from tea leaves

Tea polyphenols were extracted using pre-existing methods after slight modifications [34]. In brief, the dried green tea leaves were first defatted with n-hexane for 5 h under stirring and dried at room temperature. The dried, defatted leaves were refluxed overnight with methanol using a Soxhlet extraction unit. The obtained extract was concentrated under vacuum with a rotary evaporator at 40 °C. The concentrated tea polyphenol extract was cooled to room temperature and used for determination of TPC via the Ty-AuNP probes and FC assay (where catechol was used as a standard) [34]. The unused extract was stored at -80 °C for further use.

3. RESULTS AND DISCUSSION

3.1 Characterization of AuNPs

The synthesized AuNPs were analyzed by visual examination and UV-vis spectroscopy. The red wine color of citrate- and Ty-AuNPs was used for preliminary confirmation of AuNPs synthesis.

Journal Pre-proof

Further, the presence of a strong LSPR band between 500 to 600 nm was used to confirm the formation of AuNPs. The shape, intensity, and position of an LSPR peak are highly dependent on the shape and size of the AuNPs [32, 35]. In the first instance citrate-AuNPs of 9, 20, 28, 36, and 45 nm sizes were employed for the probe fabrication. As aforementioned, 9 nm sized AuNPs were found superior among all tested sizes in terms of sensing signal change (refer to Section 3.2). Note that all the AuNPs mentioned hereafter are 9 nm citrate-AuNPs, unless otherwise specified. In citrate-AuNPs, the LSPR band was observed at 522 nm, while the absorption peaks of Ty-AuNPs (5), Ty-AuNPs (10), Ty-AuNPs (50), and Ty-AuNPs (100) were observed at 524, 526, 530, and 535 nm, respectively **(Figure 1)**. As evident in the peak position of Ty-AuNPs, a red shift in the LSPR peak of AuNPs was observed in comparison to that of citrate-AuNPs. This shift can be attributed to Ty immobilization on citrate-AuNPs accompanied by surface modification and an associated slight increase in AuNP size [26]. To show the peak shift after Ty immobilization, we equalized the absorbance among cases.

The surface zeta potentials of citrate- and Ty-AuNPs were measured utilizing DLS. Due to the presence of free COO^- groups in citrate-AuNPs, a negative zeta potential of -45.6 mV was measured. On the other hand, immobilization of Ty provided a relatively low negative surface zeta potential for Ty-AuNPs (5), Ty-AuNPs (10), Ty-AuNPs (50), and Ty-AuNPs (100), of -44.6, -34.6, -28.7, and -17.77 mV, respectively. This change in AuNP surface can be attributed to the overall positive charge of the enzyme [36].

As observed in the TEM and SEM micrographs shown in **Figures 2 and 3**, respectively, Ty-AuNPs had spherical morphology with an average diameter of 9 nm. Elemental mapping performed via EDX showed an increasing ratio of Cu:Au in Ty-AuNPs with an increase in Ty concentration (e.g., from 5 to 100 µg/mL). The increase in Cu signal can be attributed to increased level of metalloenzyme on the AuNP surface (**Table S1).** Quantification of immobilized Ty on Ty-AuNP probes was

Journal Pre-proof

performed by BCA protein assay (**Figure S3)**. The supernatants of Ty-AuNP probes were tested for estimation of unbound enzyme in order to quantify the extent of Ty immobilization on the AuNP surface. There were 4.8, 9, 36, and 70 µg/mL of immobilized enzymes on the Ty-AuNPs (5), Ty-AuNPs (10), Ty-AuNPs (50), and Ty-AuNPs (100), respectively.

3.2 Sensing performance of the Ty-AuNP probe

The Ty-AuNP sensing probe was designed by varying the level of adsorbed Ty. The changes in LSPR peak intensity and position of the developed sensing probe were examined to quantify a diverse concentration range of catechol. The LSPR peak intensity increased with an increase in concentration of catechol. Moreover, a red shift in peak position and peak broadening also appeared after addition of catechol. The observed broadening and increased peak intensity could be related to quinone formation, whereas the shift in intensity could be attributed to NaCl-derived instability and aggregation of AuNPs. Out of all the AuNPs with different sizes (9-45 nm), AuNPs with 9 nm of size displayed the best response (33% change in the signal intensity) for the catechol. With increasing size of AuNPs, the signal response was lowered drastically (Section S4). For instance, 45 nm sized AuNPs displayed only 0.89% of the initial signals upon the addition of catechol (**Figure S4**).

After the mixing of catechol with Ty-AuNPs nanobioprobes, the LSPR peak at 530 nm was found to increase. Consequently, the absorbance at 530 nm was used for spectroscopic detection of catechol. Consequently, depending on the concentration of Ty, nanobioprobes with diverse sensitivities were obtained (**Figure 4a-d)**. Of all the fabricated sensing probes, Ty-AuNPs (5) most efficiently sensed catechol in the range of 0.01-50 ppb (**Figure 4a)**. On the other hand, Ty-AuNPs (10) performed reliably for a wide concentration range (0.01-10,000 ppm) (**Figure 4b)**. Interestingly, Ty-AuNPs (50) also worked well within a very wide range of catechol concentrations, 10-50,000 ppm, as observed in

Journal Pre-proof

Figure 4c. In addition, Ty-AuNPs (100) worked effectively for substrate concentrations of 1000-50,000 ppm. The limit of detection (LOD) values for catechol, estimated by Ty-AuNPs (5), Ty-AuNPs (10), Ty-AuNPs, and (50) Ty-AuNPs (100), were 0.00001, 0.15, 0.17, and 0.2 ppm, respectively. We detected catechol level as low as 0.01 ppb with the naked eye. As shown in **Table S2**, the developed probe is far more sensitive than the other nanomaterials-based sensing probes developed for the determination of phenolic compounds.

The sensing parameters (e.g., sensing range and LOD) of the Ty-AuNP probes were heavily dependent on the concentration of Ty used during probe fabrication. In particular, the LOD values of the sensing probe increased with increase in enzyme load. This trend in LOD can be explained by increased AuNP stability with an increase in enzyme concentration (to an extent). The color of AuNPs with higher enzyme coatings (e.g., Ty-AuNPs (50) and Ty-AuNPs (100)) was not affected by lower concentrations of catechol. For example, Ty-AuNPs (50) and Ty-AuNPs (100) could sense an upper limit of 50,000 ppm catechol, while their lower detection limits were 0.17 and 0.2 ppm, respectively. Both the above-mentioned probes were also stable after subsequent addition of NaCl or low level of catechol. It is suspected that the conformational changes induced in Ty with lower levels of catechol were not significant enough to destabilize Ty-AuNPs (100). As a result, NaCl ions were unable to interact with AuNPs or produce any color change. In contrast, AuNPs with lower Ty concentrations were susceptible and sensitive to catechol interaction. Moreover, high levels of catechol led to a blue shift in LSPR peak positions (**Figure 4a-d**). Enzyme inhibition caused by excessive substrate levels is suspected of driving the blue LSPR peak shift of Ty-AuNPs [37, 38].

It should be addressed that AuNPs functionalized with CTAB and GSH were also examined for the enzymatic sensing of catechol (refer to Section S5). In case of GSH-AuNPs, the peak at 523 nm was shifted to 539 nm after the attachment of Ty, while there was no such shift for CTAB-AuNPs after

Journal Pre-proof

the attachment of Ty. After the addition of 5000 ppm catechol, GSH-AuNPs and CTAB-AuNPs biosensing probe exhibited 28% and 0.16% of increase in the peak intensity, respectively **(Figure S5)**. Under similar sensing conditions, citrate-AuNPs based biosensing probe displayed 33% change in initial signals. In case of CTAB-AuNPs, the SPR peak was broadened significantly after the addition of catechol **(Figure S5)**. Thus, in addition to the citrate-AuNPs, GSH-AuNPs could also be a good option for the development of nanobioprobe. Furthermore, the feasibility of utilizing AgNPs was also examined **(Figure S6)**. However, as described in Section S6, due to overlapping of AgNPs SPR absorbance (398 nm) with the quinone (350-450 nm), the AgNPs-based sensing probe could not be a good choice for the current approach. Note that quinones can be produced as a product of enzymatic reaction between Ty and catechol.

3.3 Sensing mechanism of the Ty-AuNP probe

Previous studies have shown that the light absorbance of AuNPs begins to decrease as they interact with Cu [39, 40]. Likewise, Ty (a Cu-containing enzyme) exerted a significant decrease in AuNP absorbance in comparison to bare AuNPs. The interactions between AuNPs and Cu within Ty can be attributed to the decreased LSPR of AuNP. Accordingly, multiple mechanisms for the change in Ty-AuNP LSPR signals are proposed for catechol sensing. After addition of catechol to the Ty-AuNP probes, three observations were noted: (1) increase in LSPR peak intensity, (2) LSPR peak broadening, and (3) LSPR peak shift. The interaction between catechol and Ty is indicated by AuNP LSPR peak broadening or increasing intensity. After addition of catechol, its interaction with Ty affects the coordination chemistry of the enzyme (trigonal to tetragonal) by altering the oxidation state of the coordinated Cu ions from Cu(I) to Cu(II) [22, 41]. Overall, a rearrangement in Ty structure occurs, and the substrate binds the specific site [22]. This conformational change can be accounted for

Journal Pre-proof

by two types of sensing mechanisms. First, as Ty interacts with catechol, the distance between AuNPs and Cu increases, and AuNP absorbance is regained. This phenomenon became more significant as catechol level increased because more Cu ions are displaced from the AuNPs. Second, AuNP destabilization caused by interaction with NaCl ions is accompanied by aggregation of AuNPs and a red shift in LSPR peak. In the presence of higher concentrations of NaCl (e.g., 0.5 M), AuNP aggregation caused change in color from red to blue [42]. The conformational changes in immobilized Ty expose the citrate-AuNP surface to interact with the NaCl ions. The presence of salt cations causes electrostatic screening of negative charges present on AuNPs [42]. As the lack of surface charges on AuNPs causes aggregation, their color is converted from red wine to blue [42]. Overall, based upon the testing results of the developed probe for catechol, two mechanisms were recognized for target sensing: the first mechanism was responsible for the increase in AuNP LSPR peak, while the second mechanism was responsible for the LSPR peak shift.

3.4 Detection of polyphenols in real samples

The developed system was tested to measure TPC level against real samples of green tea and tap water to validate the feasibility of the designed Ty-AuNPs probes. Due to the broad and suitable detection range, Ty-AuNP (50) probes were selected for TPC detection in green tea and groundwater samples. As shown in **Figure 5**, around 450 ppm of extracted tea polyphenols was detected using the Ty-AuNP (50) sensing probe. In contrast, when the FC method was applied to the same extract, it was able to detect roughly 400 ppm of polyphenols. The developed probe was also examined to detect TPC directly on tea leaves extract. By targeting the tea leaf extract, a lengthy TPC extraction step can be omitted. Interestingly, the Ty-AuNPs (50) were equally effective at determining the TPC levels in raw leaves (i.e., equivalent to the levels obtained for

Journal Pre-proof

TPC extracts). Therefore, the developed sensor shows highly selective signals for TPC, even in a complex medium. In contrast, the TPC levels in untreated leaf extracts were overestimated by a factor of two, relative to TPC extract samples when the FC method was used. The FC results of the latter assay should be more accurate as the effects of interfering molecules (e.g., vitamins and fats) can be removed via solvent extraction [43]. Additionally, the developed sensor displayed a recovery of 96-101% in tap water samples spiked with 10,000 ppm of catechol. In contrast, the FC method recorded a recovery of 87-92% for the same samples. The developed Ty-AuNP-based sensing probe is more sensitive, selective, and simple in comparison to the common alternatives. Overall, this dual-purpose probe can be used for quantitative analysis of useful and harmful phenols, such as those found in plants and tap water, respectively.

3.5 Specificity study of the Ty-AuNP probe

As discussed above, the FC method was more likely to overestimate the TPC level in untreated tea leaf extract. Such bias might reflect the possible interaction of FC reagents with other components, such as vitamins, specifically ascorbic acid. Consequently, sensing experiments were performed to estimate the levels of catechol in the presence of interfering agents such as ascorbic acid, acetic acid, and formic acid. As shown in **Figure S7,** none of interfering agents affected the LSPR of the designed probe, whereas the traditional method showed change in color in the presence of ascorbic acid. Moreover, the probe was efficient enough to detect the catechol in the presence of mixture of all the above mentioned interfering agents (**Figure S7**). On the contrary, the FC probe displayed false-positive results in the presence of ascorbic acid.

Journal Pre-proof

3.6 Reusability

To analyze the reusability of the developed sensing probe, we separated the sensing probe from the analyte solution using a dialysis membrane tube after the first sensing experiment and in the absence of NaCl. The majority of Ty-AuNPs was retained in the dialysis assembly, while the analyte or bi-products of the analyte were collected as filtrate. Interestingly, after removal of substrate and associated product, the absorbance of Ty-AuNPs decreased, indicating an interaction between the Cu present in Ty and the AuNPs. After addition of fresh substrate, catechol interacted with retained Ty-AuNPs and increased the SPR signal of the Ty-AuNP probe (**Figure 5)**. Importantly, the percentage increase in Ty-AuNP peak intensity for the second sensing cycle was similar to that of the first cycle before dialysis compared with their control (e.g., Ty-AuNPs before the addition of analyte for the respective cycle) measurements. After the fifth cycle of sensing, a 15% decrease in signals (the percentage increase in Ty-AuNPs peak intensity) was observed in comparison to the first cycle. Hence, the designed probe can be used for multiple analyses of phenolic compounds, adding an advantage to Ty-AuNP probes in comparison to other analytical methods or techniques.

Journal Pre-proof

Conclusion

Herein, we describe a sensitive, selective, and straightforward Ty-AuNP sensing probe (spectroscopic and color-based) for detection of TPC. Due to its simplicity, this probe can be used as a viable option for practical applications (e.g., grading of food components and examining water contamination). By varying the Ty level in AuNPs, a wide range (0.00001-50,000 ppm) of TPC can be determined. On the basis of diverse sensing ranges, the current probe can be used as a common tool for identification of useful phenols (e.g., polyphenols) and harmful phenolic pollutants. We also identified

and proposed a unique sensing phenol mechanism (depending upon the interactions between Cu-containing enzyme and AuNPs) of our Ty-AuNP probe. The developed Ty-AuNP sensing probe showed an LOD of 0.01 ppb for catechol. Moreover, using the Ty-AuNP sensing probe, TPC in tea and water samples were examined and quantified successfully.

Acknowledgment

This research acknowledges support from the Department of Science and Technology, New Delhi, India, in the form of an INSPIRE Faculty Award (DST/INSPIRE/04/2017/002953) and from the Science and Engineering Research Board (SERB), Government of India, under the Early Career Research (ECR) award (File No. ECR/2018/000748). KHK acknowledges support by a grant from the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ITC (MSIT) of the Korean government (Grant No: 2021R1A3B1068304).

Journal Pre-proof

References

[1] R.-Y. Gan, C.-L. Chan, Q.-Q. Yang, H.-B. Li, D. Zhang, Y.-Y. Ge, et al., 9 - Bioactive compounds and beneficial functions of sprouted grains, in: H. Feng, B. Nemzer, J.W. DeVries (Eds.), Sprouted Grains, AACC International Press2019, pp. 191-246.
[2] E.P. Meulenberg, Phenolics: occurrence and immunochemical detection in environment and food, Molecules, 14(2009) 439-73.
[3] H. Cory, S. Passarelli, J. Szeto, M. Tamez, J. Mattei, The Role of Polyphenols in Human Health and Food Systems: A Mini-Review, Front Nutr., 5(2018) 87-.
[4] D.J. Bhuyan, A. Basu, Phenolic compounds: Potential Health Benefits and Toxicity, 2017, pp. 27-59.
[5] J. Karam, M.D.M. Bibiloni, J.A. Tur, Polyphenol estimated intake and dietary sources among older adults from Mallorca Island, PLoS One, 13(2018) e0191573-e.
[6] C. Del Bo, S. Bernardi, M. Marino, M. Porrini, M. Tucci, S. Guglielmetti, et al., Systematic Review on Polyphenol Intake and Health Outcomes: Is there Sufficient Evidence to Define a Health-Promoting Polyphenol-Rich Dietary Pattern?, Nutrients, 11(2019) 1355.
[7] R. Álvarez, H. Araya, R. Navarro-Lisboa, C. Lopez de Dicastillo, Evaluation of Polyphenol Content and Antioxidant Capacity of Fruits and Vegetables Using a Modified Enzymatic Extraction, Food Technol. Biotechnol., 54(2016) 462-7.

[8] M. Kopjar, M. Tadić, V. Piližota, Phenol content and antioxidant activity of green, yellow and black tea leaves, Chem. Biol. Technol. Agric., 2(2015) 1.
[9] J.-Y. Lin, C.-Y. Tang, Determination of total phenolic and flavonoid contents in selected fruits and vegetables, as well as their stimulatory effects on mouse splenocyte proliferation, Food Chem., 101(2007) 140-7.
[10] I. Ignat, I. Volf, V.I. Popa, A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables, Food Chem., 126(2011) 1821-35.
[11] L.G. Malta, R.H. Liu, Analyses of Total Phenolics, Total Flavonoids, and Total Antioxidant Activities in Foods and Dietary Supplements, in: N.K. Van Alfen (Ed.) Encyclopedia of Agriculture and Food Systems, Academic Press, Oxford, 2014, pp. 305-14.
[12] L.G.C. Villegas, N. Mashhadi, M. Chen, D. Mukherjee, K.E. Taylor, N. Biswas, A Short Review of Techniques for Phenol Removal from Wastewater, Curr. Pollut. Rep., 2(2016) 157-67.
[13] M.J. Brandt, K.M. Johnson, A.J. Elphinston, D.D. Ratnayaka, Chapter 7 - Chemistry, Microbiology and Biology of Water, in: M.J. Brandt, K.M. Johnson, A.J. Elphinston, D.D. Ratnayaka (Eds.), Twort's Water Supply (Seventh Edition), Butterworth-Heinemann, Boston, 2017, pp. 235-321.
[14] R. Noori, M. Perwez, J.A. Mazumder, M. Sardar, Development of low-cost paper-based biosensor of polyphenol oxidase for detection of phenolic contaminants in water and clinical samples, Environ. Sci. Pollut. Res., 27(2020) 30081-92.
[15] W.W. Anku, M. Mamo, P. Govender, Phenolic Compounds in Water: Sources, Reactivity, Toxicity and Treatment Methods, 2017, pp. 420-43.
[16] J.C. Carmona-Hernandez, G. Taborda-Ocampo, C.H. González-Correa, Folin-Ciocalteu Reaction Alternatives for Higher Polyphenol Quantitation in Colombian Passion Fruits, International Journal of Food Science, 2021(2021) 8871301.
[17] J.D. Everette, Q.M. Bryant, A.M. Green, Y.A. Abbey, G.W. Wangila, R.B. Walker, Thorough study of reactivity of various compound classes toward the Folin-Ciocalteu reagent, J. Agric. Food Chem., 58(2010) 8139-44.
[18] M. Way, J. Jones, D. Nichols, R. Dambergs, N. Swarts, A Comparison of Laboratory Analysis Methods for Total Phenolic Content of Cider, Beverages, 6(2020) 55.
[19] C. Erkmen, S. Kurbanoglu, B. Uslu, Fabrication of poly(3,4-ethylenedioxythiophene)-iridium oxide nanocomposite based Tyrosinase biosensor for the dual detection of catechol and azinphos methyl, Sens. Actuators, B, 316(2020) 128121.
[20] F. Campanhã Vicentini, L.L.C. Garcia, L.C.S. Figueiredo-Filho, B.C. Janegitz, O. Fatibello-Filho, A biosensor based on gold nanoparticles, dihexadecylphosphate, and tyrosinase for the determination of catechol in natural water, Enzyme Microb. Technol., 84(2016) 17-23.
[21] I.M. Apetrei, M.L. Rodriguez-Mendez, C. Apetrei, J.A. de Saja, Enzyme sensor based on carbon nanotubes/cobalt(II) phthalocyanine and tyrosinase used in pharmaceutical analysis, Sens. Actuators, B, 177(2013) 138-44.
[22] Y. Matoba, S. Kihara, N. Bando, H. Yoshitsu, M. Sakaguchi, K.e. Kayama, et al., Catalytic mechanism of the tyrosinase reaction toward the Tyr98 residue in the caddie protein, PLoS Biol, 16(2018) e3000077-e.
[23] L. Scotti, F.J.B.M. Júnior, H.M. Ishiki, F.F. Ribeiro, M.C. Duarte, G.S. Santana, et al., Chapter 9 - Computer-Aided Drug Design Studies in Food Chemistry, in: A.M. Grumezescu, A.M. Holban (Eds.), Natural and Artificial Flavoring Agents and Food Dyes, Academic Press2018, pp. 261-97.
[24] F.M. Fartas, J. Abdullah, N.A. Yusof, Y. Sulaiman, M.I. Saiman, Biosensor Based on Tyrosinase Immobilized on Graphene-Decorated Gold Nanoparticle/Chitosan for Phenolic Detection in Aqueous, Sensors (Basel), 17(2017) 1132.

Journal Pre-proof

[25] V. Carralero Sanz, M.L. Mena, A. González-Cortés, P. Yáñez-Sedeño, J.M. Pingarrón, Development of a tyrosinase biosensor based on gold nanoparticles-modified glassy carbon electrodes: Application to the measurement of a bioelectrochemical polyphenols index in wines, Anal. Chim. Acta, 528(2005) 1-8.
[26] M.F. Siddiqui, S. Jeon, M.-M. Kim, Monitoring of whitening agent for skin analysis using tyrosinase gold nanoparticle-based colorimetric assay, Asia-Pac. J. Chem. Eng., 16(2021) e2593.
[27] D.K. Nguyen, C.-H. Jang, Ultrasensitive colorimetric detection of amoxicillin based on Tris-HCl-induced aggregation of gold nanoparticles, Anal. Biochem., 645(2022) 114634.
[28] T. Patil, R. Gambhir, A. Vibhute, A.P. Tiwari, Gold Nanoparticles: Synthesis Methods, Functionalization and Biological Applications, J. Cluster Sci., (2022).
[29] O. Adegoke, S. Zolotovskaya, A. Abdolvand, N.N. Daeid, Fabrication of a near-infrared fluorescence-emitting SiO2-AuZnFeSeS quantum dots-molecularly imprinted polymer nanocomposite for the ultrasensitive fluorescence detection of levamisole, Colloids Surf., A, 646(2022) 129013.
[30] O. Adegoke, N.N. Daeid, Polymeric-coated Fe-doped ceria/gold hybrid nanocomposite as an aptasensor for the catalytic enhanced colorimetric detection of 2,4-dinitrophenol, Colloids Surf., A, 627(2021) 127194.
[31] O. Adegoke, S. Zolotovskaya, A. Abdolvand, N.N. Daeid, Rapid and highly selective colorimetric detection of nitrite based on the catalytic-enhanced reaction of mimetic Au nanoparticle-CeO2 nanoparticle-graphene oxide hybrid nanozyme, Talanta, 224(2021) 121875.
[32] K. Vaid, J. Dhiman, N. Sarawagi, V. Kumar, Experimental and computational study on the selective interaction of functionalized gold nanoparticles with metal ions: Sensing prospects, Langmuir, 36(2020) 12319-26.
[33] M.Y. Marcelino, F.A. Borges, L. Scorzoni, J. de Lacorte Singulani, B.C. Garms, J.C. Niemeyer, et al., Synthesis and characterization of gold nanoparticles and their toxicity in alternative methods to the use of mammals, J. Environ. Chem. Eng., 9(2021) 106779.
[34] P. Khare, R. Maurya, R. Bhatia, P. Mangal, J. Singh, K. Podili, et al., Polyphenol rich extracts of finger millet and kodo millet ameliorate high fat diet-induced metabolic alterations, Food Funct., 11(2020) 9833-47.
[35] M. Pannico, A. Calarco, G. Peluso, P. Musto, Functionalized Gold Nanoparticles as Biosensors for Monitoring Cellular Uptake and Localization in Normal and Tumor Prostatic Cells, Biosensors (Basel), 8(2018) 87.
[36] C. Garcia-Hernandez, A.K. Freese, M.L. Rodriguez-Mendez, A.K. Wanekaya, In situ synthesis, stabilization and activity of protein-modified gold nanoparticles for biological applications, Biomater. Sci., 7(2019) 2511-9.
[37] M. Yoshino, K. Murakami, Analysis of the substrate inhibition of complete and partial types, Springerplus, 4(2015) 292-.
[38] R.R. Ramsay, K.F. Tipton, Assessment of Enzyme Inhibition: A Review with Examples from the Development of Monoamine Oxidase and Cholinesterase Inhibitory Drugs, Molecules (Basel, Switzerland), 22(2017) 1192.
[39] W. Yan, Z. Zhong, J. Ma, T. Rujiralai, Highly sensitive colorimetric sensing of copper(ii) ions based on “CLICK-17” DNAzyme-catalyzed azide modified gold nanoparticles and alkyne capped dsDNA cycloaddition, RSC Adv., 11(2021) 24196-205.
[40] Z. Weng, H. Wang, J. Vongsvivut, R. Li, A.M. Glushenkov, J. He, et al., Self-assembly of core-satellite gold nanoparticles for colorimetric detection of copper ions, Anal. Chim. Acta, 803(2013) 128-34.
[41] T. Mulla, S. Patil, S. Sistla, J. Jadhav, The Binding Affinity of Small Molecules with Yam Tyrosinase (Catechol Oxidase): A Biophysical Study, Biochem. Res. Int., 2019(2019) 8284968.
[42] S. Christau, T. Moeller, J. Genzer, R. Koehler, R. von Klitzing, Salt-Induced Aggregation of Negatively Charged Gold Nanoparticles Confined in a Polymer Brush Matrix, Macromolecules, 50(2017) 7333-43.

Journal Pre-proof

[43] S.k. Abdul mudalip, I. Izhab, S. Abdullah, Z. Iffah, W. Jalil, Extraction of Vitamin C from Red Pitaya via Soxhlet Extraction Method2009.

Author Biographies

Ms. Jasmeen Dhiman

Ms. Jasmeen Dhiman was research fellow at National Agri-Food Biotechnology Institute (NABI), Mohali, India. She holds a Master of Science (Hons.) degree in the stream of Microbial and Food Technology, from Punjabi University, Patiala, India. Her current area of research includes development of nanomaterial based sensors/ biosensors for the detection of heavy metals and pesticides in food and water samples.

Mr. Kalyan Vaid

Mr. Kalyan Vaid is presently working as ICMR-SRF at CSIR-CSIO, Chandigarh, India. Before joining CSIR-CSIO, he was SRF at the National Agri-Food Biotechnology Institute (NABI), Mohali, India. He is academically trained in the field of Biophysics from Panjab University, Chandigarh. His main area of research includes development of nanomaterials based electrochemical and optical sensors for heavy metals and pesticides in food and water.

Ms. Treesa Johns

Treesa Johns received MSc degree in Analytical chemistry from Mahatma Gandhi University Kottayam, Kerala. She is currently working as a Project associate at Hindustan petroleum Green R and D center (HPGRDC), Bengaluru. She did her student internship from National Agri-Food Biotechnology Institute (NABI), Mohali, India. Her research interest includes development of biosensors for food contaminants, electrochemistry, and application of nanotechnology in Li-ion batteries.

Journal Pre-proof

Ms. Ruchika Maurya

Ruchika Maurya is currently working as a PhD fellow under NABI-RCB-PhD program at the National Agri-Food Biotechnology Institute (NABI), Mohali, India. She has achieved her M.Sc. degree in Life Science from Central University of Punjab. Her current area of research includes cereal bioactive ingredients in ameliorating metabolic syndrome.

Ms. Mahima Arora

Mahima Arora received her master degree in Biotechnology from Panjab University, Chandigarh. She is currently working as Business Development Analyst at Signicent LLP, Mohali, India. She did her thesis from National Agri-Food Biotechnology Institute (NABI), Mohali, India. Her research interest includes development of nano-biosensors for the detection of contaminants in food and development of nano cellulosic fibres from waste materials.

Ms. Ankita Negi

Ankita Negi is currently working as a Project Fellow at National Agri-Food Biotechnology Institute (NABI), Mohali, India. She has completed her master degree in Biotechnology from Delhi Technological University (DTU), New Delhi, India. Her current area of research includes development of nanomaterial based sensors/ biosensors for the detection of heavy metals and pesticides in food and water samples.

Ms. Ritika Gupta

Ms. Ritika Gupta has received her post-graduate degree in biotechnology from Banasthali Vidyapeeth, Rajasthan. She has previous experience in biosensing of two years from CSIR-Institute of Microbial Technology Institute, Chandigarh, India. She is currently a Ph.D. research scholar at National Agri-Food Biotechnology Institute, Mohali, India. Her research interests include designing aptamers against whole-cell bacteria and proteins. Further, she is also developing various biosensing assays for the fast detection of bacteria using nanomaterials. She

Journal Pre-proof

has research papers in reputed journals like Sensors and Actuators B: Chemical, Microchimica Acta, ACS applied materials and interfaces, and Nature scientific reports.

Dr. Mrinmoy Misra

Dr. Mrinmoy Misra is an Assistant Professor in the Department of Mechatronics Engineering, Manipal University Jaipur, India. He received his PhD from the Academy of Scientific and Innovative Research (AcSIR), in September 2015. After being a postdoctoral researcher in Indian Institute of Technology Kanpur, India (April 2015 – March 2016) and Research Institute of Battery and Energy Conservation, Gachon University, South Korea (Sep. 2016 to March 2018) and as an assistant professor in Gachon University, South Korea (April 2018 to March 2020), he joined as an assistant professor, Manipal University Jaipur, India. His main research areas are transparent conducting thin film heaters fabrication and nanomaterial-based energy generation and core-shell nanomaterial synthesis and application.

HYU Distinguished Professor Ki-Hyun Kim

HYU Distinguished Prof. Ki-Hyun Kim completed an MS at Florida State University (1986) and a PhD at University of South Florida (1992). In 2014, he joined Department of Civil & Environmental Engineering at Hanyang University. His research interfaces Environmental and Material Engineering through Coordination Polymers. He became one of the top **10 National Star Faculties in Korea in 2006** and an **academician** (Korean Academy of Science and Technology) in 2018. He has been recognized as **'Highly Cited Researcher (HCR)' in 'Environment & Ecology' field as of 2019 and 'Engineering' field as well from 2021**. He has published more than 890 articles in leading scientific journals including '**Chemical Society Reviews'.**

Dr. Vanish Kumar

Dr Vanish Kumar is presently working as an Assistant Professor at Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala, India. Before joining TIET, Patiala, he was INSPIRE faculty at the National Agri-Food Biotechnology Institute (NABI), Mohali,

Journal Pre-proof

India. He has an M. Tech degree in nanoscience and nanotechnology. He completed his PhD (Engineering) from Academy of Scientific and Innovative Research (AcSIR) CSIR-CSIO, Chandigarh, India. His current areas of research include development of advanced nanostructures for sensing/removal of contaminants/pollutants and analytical applications of nanomaterials. He has published more than 80 research papers in reputed international journals.

Figures

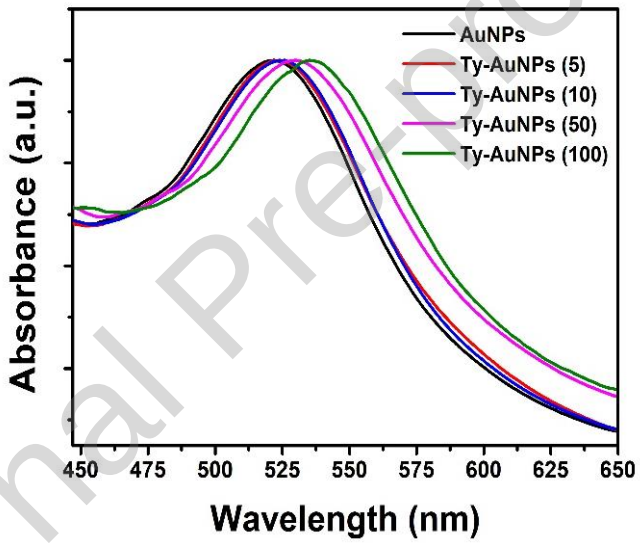
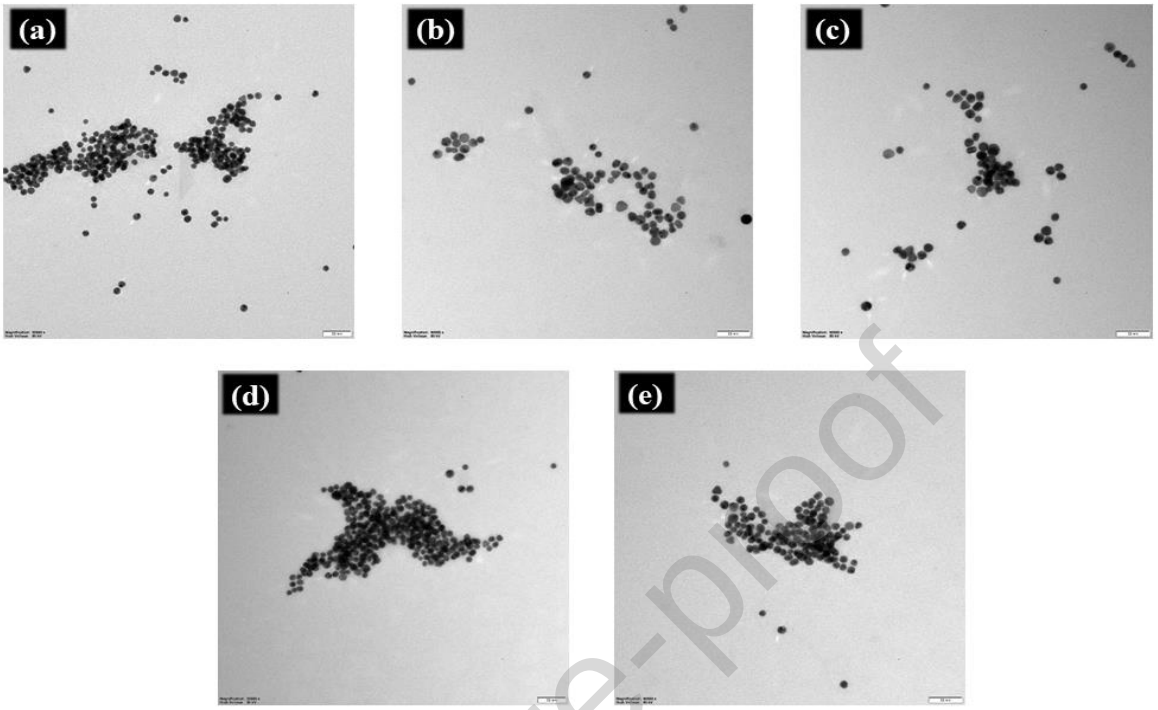


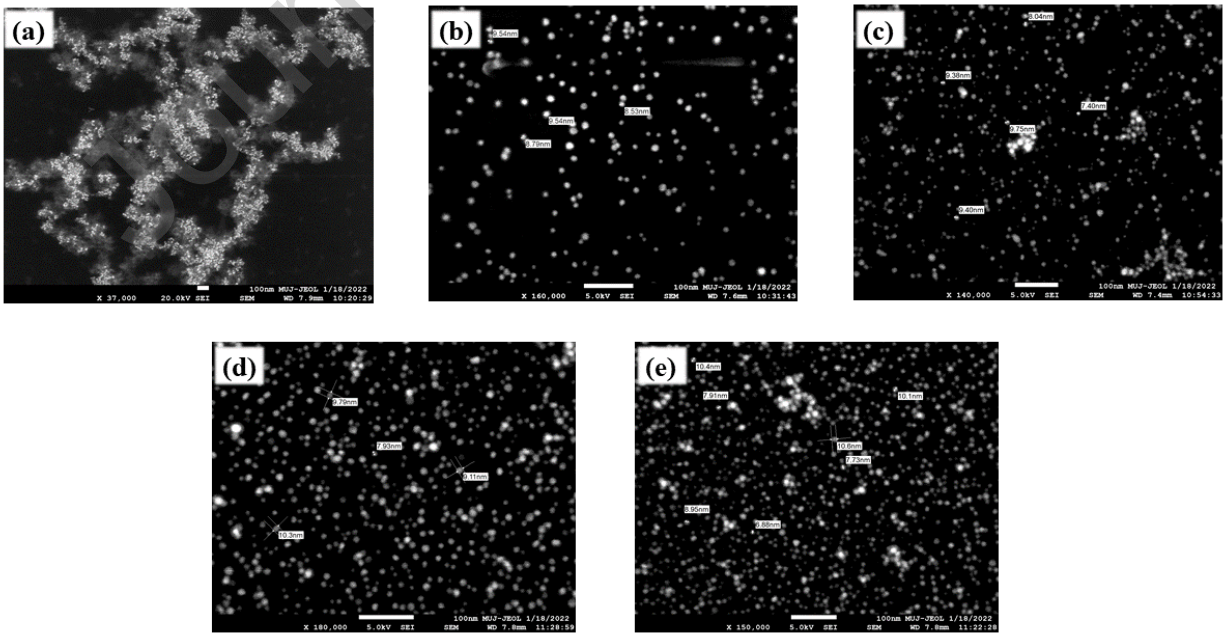
Figure 1. UV-Vis absorption spectrum of synthesized citrate-AuNPs and Ty-immobilized AuNPs (Ty-AuNPs) with different enzyme concentrations (5, 10, 50, and 100 µg/mL).

Journal Pre-proof



.

Figure 2. TEM images of AuNPs and AuNPs-based sensing probes: (a) Citrate-AuNPs, (b) Ty-AuNPs (5), (c) Ty-AuNPs (10), (d) Ty-AuNPs (50), and (e) Ty-AuNPs (100).



Journal Pre-proof

Figure 3. SEM micrograms of AuNPs and AuNPs-based sensing probes: (a) Citrate-AuNPs, (b) Ty-AuNPs (5), (c) Ty-AuNPs (10), (d) Ty-AuNPs (50), and (e) Ty-AuNPs (100).

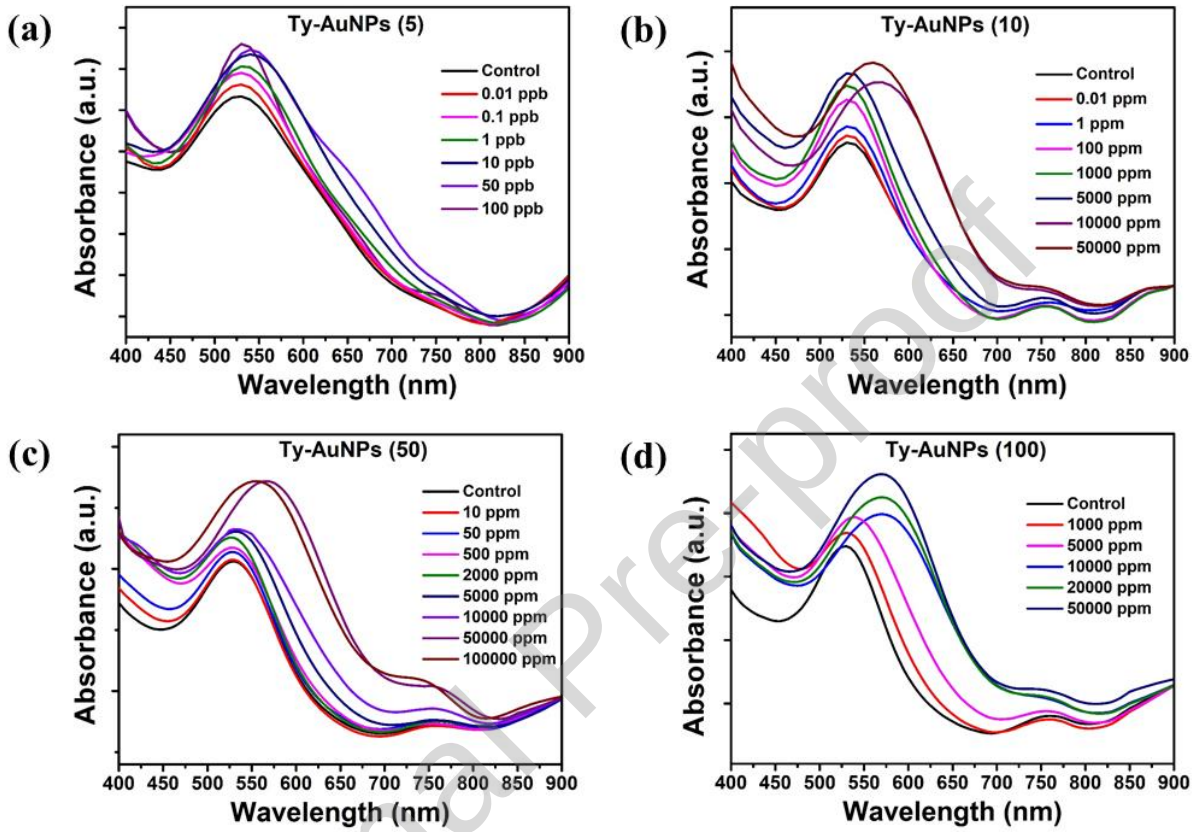
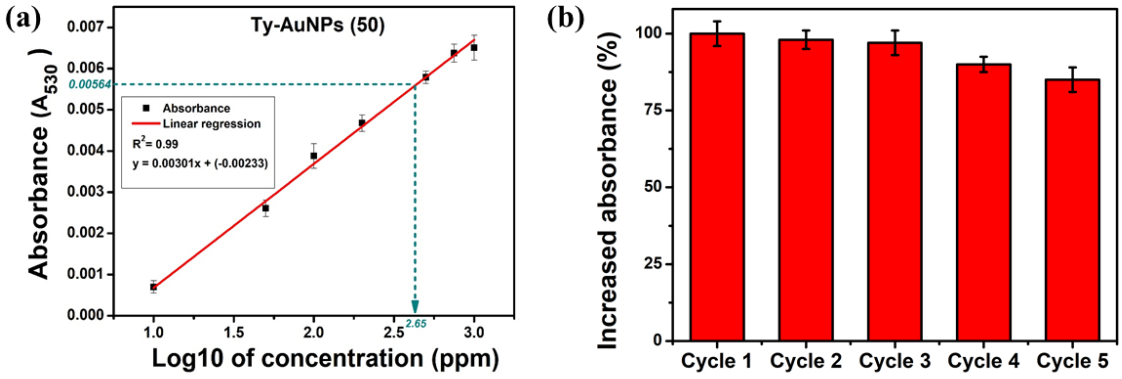


Figure 4. UV-vis spectra showing the effects of catechol concentration on the absorbance of Ty-AuNPs sensing probe: (a) Ty-AuNPs (5), (b) Ty-AuNPs (10), (c) Ty-AuNPs (50), and Ty-AuNPs (100).



Journal Pre-proof

Figure 5. Testing of Ty-AuNPs (50) for detection of TPC in tea and its reusability potential: (a) concentration of catechol vs. absorbance of Ty-AuNPs (50) and standard curve for determination of TPC level in tea leaves. (b) Bar graph showing the reusability of Ty-AuNPs (50) for determination of TPC.

CRediT authorship contribution statement

Jasmeen Dhiman: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, review & editing; **Kalyan Vaid:** Data curation, Investigation, Formal analysis, Methodology, Visualization, Writing: review & editing; **Treesa Johns;** Conceptualization, Formal analysis, Investigation, Writing: review & editing; **Ruchika Maurya:** Methodology, Formal analysis, Investigation, Writing: review & editing; **Mahima Arora:** Formal analysis, Investigation, Writing: review & editing; **Ankita Negi** Methodology, Formal analysis, Investigation, Writing: review & editing; **Ritika Gupta:** Formal analysis, Investigation, Writing: review & editing; **Mrinmoy Misra; :** Formal analysis, Investigation, Supervision, Writing: review & editing; **Ki-Hyun Kim:** Formal analysis, Funding acquisition, Supervision, Validation, Writing- review & editing; **Vanish Kumar:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing- original draft, review & editing. All authors have contributed to manuscript reviewing and editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Highlights

- Ty-AuNPs probes were developed for ultrasensitive sensing of phenolic compounds.
- The spectroscopic and color-based sensing was performed.
- The sensor works well for 0.01 ppb- 50,000 ppm of phenolic compounds.
- The lowest LOD value of 0.01 ppb was achieved by Ty-AuNPs (5).
- A novel Cu-based sensing mechanism has been proposed.

Journal Pre-proof

Voltage Regulation in a Radial Microgrid with High RES Penetration

Approach-Optimum DVR Control

Ritika Gour
Department of Electrical Engineering
Delhi Technological University
Delhi, India
riti113@gmail.com

Vishal Verma
Department of Electrical Engineering
Delhi Technological University
Delhi, India
vishalverma@dce.ac.in

Received: 4 April 2022 | Revised: 26 April 2022 and 9 May 2022 | Accepted: 12 May 2022

Abstract-Increasing installation of Renewable Energy Sources (RESs) in microgrids and distribution system feeders benefits the power system in many ways, such as increasing power availability and reliability, but they also cause various power quality issues, such as overvoltage and undervoltage, due to their intermittent nature. The sensitive loads, such as Constant Power Load (CPL), are affected by these voltage fluctuations. The approach proposed in this paper is to install a Dynamic Voltage Restorer (DVR) adjacent to the RES to suppress the voltage issue caused by it while leaving the sensitive load operation unharmed. The proposed concept is duly endorsed with analysis and simulations in MATLAB/Simulink environment.

Keywords-DVR; intermittent RES; CPL; voltage regulation; microgrid

I. INTRODUCTION

Penetration of Renewable Energy Sources (RESs), such as solar, wind, hydro energy, etc. has been increased in the distribution systems due to increased demand, availability, and the environmental concerns caused by the conventional sources [1]. When the penetration is increased, a part of the distribution feeder needs to be controlled and monitored separately by the distributed system operator. This part of the feeder may be regarded as a microgrid [2-3]. The increased penetration of RESs and their independent control gives to the system many advantages, such as increased available power, lesser generation from conventional power sources, increased power reliability, etc.. However, the increased penetration of RESs in the feeder may give rise to some power quality issues, in which variation in voltage level due to the intermittent nature of the RES is of major concern, especially today, when many of the loads connected in the distribution system require constant power for their better operation and behave like Constant Power Loads (CPLs) [4-7]. The power consumed by a CPL remains constant regardless of the voltage at the point of common coupling. If the voltage drops, the current level rises and vice versa. When tightly regulated, power electronic converters such as dc/dc converters feeding a load such as a battery, or dc/ac inverters driven loads are presumed to be CPLs. Furthermore, switching power supplies, which are found in the majority of mobile systems and inverter-based motor drive systems, act in the same way as CPLs [3, 5-6].

The power requirement of these CPL loads is unaffected by voltage variations at the terminal. When the voltage at the CPL terminals is increased, the current demanded by these loads decreases, resulting in negative impedance characteristics as observed by the feeder. This type of CPL characteristic could exacerbate the voltage rise issue in the feeder. As the current requested by the CPL drops, the overall drop in the feeder lowers as well, resulting in a higher voltage at the CPL's terminal and greater current reversal towards the main grid source. These factors may reduce the voltage stability of the microgrid. If the number of such CPLs connected in the micro-grid increases, a cascading effect could occur, causing instability in the micro-grid/distribution feeder [8]. To minimize such issues, the voltage at the CPL terminals should be kept within tight restrictions, and be such that the RES intermittency has the least impact on CPLs. There are numerous devices reported in the literature that can be used to regulate voltage in the distribution system, including Dynamic Voltage Restorers (DVRs), Static Synchronous Compensators (STATCOMs), etc. [9-12]. A DVR is a custom power device that is connected in series with the line and injects/absorbs a series voltage. In this paper, a radial microgrid voltage regulation approach is presented, in which the DVR is situated near the RES, minimizing the issue produced on the CPL connected in the radial microgrid by the intermittent nature of the RES. The voltage at the DVR's terminals is regulated by injecting/absorbing a voltage in series with the feeder. The proposed method is backed up by analysis and is tested in Matlab/Simulink environment.

II. SYSTEM CONFIGURATION

Figure 1 depicts the system diagram for the grid-connected radial microgrid with CPL and intermittent RES. The RES unit comprised of an RES source such as a rooftop PV or a wind turbine and is connected to the micro-grid/distribution system via a Voltage Source Converter (VSC) that is connected to the feeder as a current source. The RES is placed very close to the

Corresponding author: Ritika Gour

CPL in order to witness and study the worst effects of RES intermittency. The DVR is installed adjacent to the RES, so that it can absorb extra power and provide a small percentage of power, if needed, to regulate the voltage. Batteries can be used to extend the DVR's regulatory and compensation limits [13]. The grid is represented by the source V_S, while the feeder impedance is represented by the Z_S. The feeder in consideration is a distribution type feeder with an R/X ratio of around 8 [14].

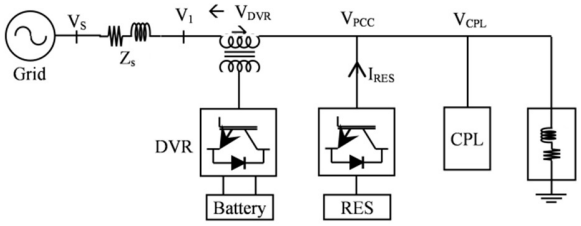


Fig. 1. Single line diagram of the considered system.

III. ANALYSIS OF THE CONSIDERED SYSTEM

The analysis of the considered microgrid system is conducted in two stages: At first, the impact of the RES on the CPL is investigated, and in the second stage, the way the existence of DVR may ameliorate the situation is investigated. The RES is considered to be situated very close to the CPL, and thus the feeder impedance between them can be overlooked.

A. Impact of RES on CPL

The equivalent circuit diagram for the considered system with only RES and CPL is shown in Figure 2(a). The Voltage at the terminals of CPL can be calculated with the help of superposition theorem and is given by:

$$V_{CPL} = \frac{Z_{CPL}(V_S + I_{RES}Z_S)}{Z_S + Z_{CPL}} \quad (1)$$

The impedance of the constant power load is dynamic in nature and can be written as:

$$Z_{CPL} = \frac{V_{CPL}^2}{P_{CPL}} \quad (2)$$

Equation (1) rewritten and simplified with value of Z_{CPL} obtained from (2) gives:

$$V_{CPL}^2 - V_{CPL}(V_S + I_{RES}Z_S) + Z_S P_{CPL} = 0 \quad (3)$$

If the RES current is known, then for a particular CPL load the voltage at the terminals of CPL can be estimated by solving (3):

$$V_{CPL} = \frac{(V_S + I_{RES}Z_S) \pm \sqrt{(V_S + I_{RES}Z_S)^2 - 4Z_S P_{CPL}}}{2} \quad (4)$$

B. With DVR Compensation Unit

In this stage the DVR is connected to the system as shown in Figure 2(b). The voltage at the terminal of RES can be regulated closed to the rated value through DVR. The voltage at the terminal of CPL obtained from (3) is taken into consideration for the estimation of the voltage to be supplied/absorbed by the DVR. The new voltage at the terminal of CPL after the insertion of DVR can also be estimated with the help of superposition theorem and is given by:

$$V_{CPL(New)} = \frac{Z_{CPL}(V_S + I_{RES}Z_S + V_{DVR})}{Z_S + Z_{CPL}} \quad (5)$$

After putting the value of Z_{CPL}, (5) can be rewritten as:

$$P_{CPL}Z_S + V_{PCC}^2 = V_{PCC}(V_S + I_{RES}Z_S + V_{DVR}) \quad (6)$$

If the voltage injected by the DVR is known, then the voltage at the terminals of the CPL can be calculated by solving the above quadratic equation. This equation can also be used in another way, for the estimation of voltage to injected/absorbed by the DVR as mentioned in (7).

$$V_{DVR} = \frac{P_{CPL}Z_S + V_{CPL}^2}{V_{CPL}} - V_S + I_{RES}Z_S \quad (7)$$

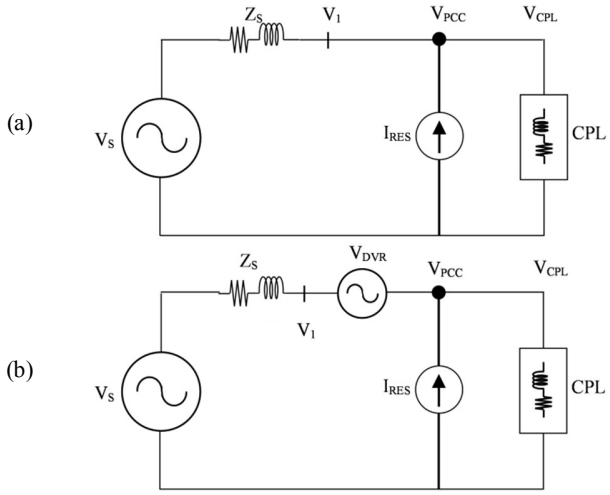


Fig. 2. Equivalent circuit diagram of the considered system (a) without any regulation from the DVR, (b) with DVR.

IV. CONTROL OF DVR

The control strategy for generating gating pulses of the voltage source inverter of the DVR is depicted in Figure 3. The synchronous reference frame theory is being used to control the DVR. The DVR injects or absorbs voltage in series to regulate the voltage at its terminal. It accomplishes this by sensing the voltage at the terminal, comparing it to the reference value of the voltage, and passing it to the Proportional Integral (PI) controller, which generates a voltage reference value for the quadrature axis (q-axis) component of the DVR. Initially, the DVR tries to keep the voltage underneath the range by using only the reactive component, resulting in only the q-axis component (V_q). However, if V_q reaches its maximum value, the q-axis component is reduced and the d-axis component is increased, resulting in a signal that may be used to estimate the voltage required to regulate the voltage with both d-axis (V_d) and q-axis (V_q) component at the DVR's terminal. To estimate the reference signal for the DVR's VSI, the reference d-axis and q-axis components are transformed into reference abc components via reverse parks transformation. To generate the gating pulse for the VSI switches, the reference signal is now routed through the pulse width modulator. The VSC then uses the obtained gate pulses to operate the switches.

V. PERFORMANCE EVALUATION OF THE PROPOSED SYSTEM

To validate the proposed concept, the considered system is simulated in MATLAB/Simulink. The simulation diagram of the system is shown in Figure 4. The system parameters used in the simulations are mentioned in Table I. To demonstrate the applicability of the proposed design, the system is tested under two conditions: (i) RES intermittency and (ii) load perturbation. For the sake of evaluation convenience, the initial transients that emerge in the simulation are ignored.

TABLE I. SIMULATION PARAMETER USED

Parameter	Parameter Values
Source/Grid Voltage	415V line to line
Feeder impedance	0.642+j0.083Ω/km [14], feeder length =0.5km
RES rating	10kW
CPL rating	3.75kW
Constant impedance load	2.5kVA to 7.5kVA, 0.8 pf lagging load
DVR rating	2kVA, maximum injection upto 100V

Figure 5 depicts the influence of RES intermittency on the voltage level in the microgrid and the ability of DVR to regulate the voltage during these intermittencies. RES injects a current of $I_{RESpeak}$ = 10A from t = 0 to 0.5s. Since the DVR is not injecting/absorbing voltage from t = 0 to 0.25s, the voltage observed (V_{CPL}) is lower than the rated value. The DVR is switched into the circuit at 0.25s, and the voltage at the DVR terminal is regulated to the rated value by injecting voltage (V_{DVR} = 71V) into the microgrid. Since RES penetration is low during this period, the DVR can regulate the voltage with only reactive power. The DVR is switched off at 0.5s, and the RES increased to more than double the initial value ($I_{RESpeak}$ = 20A) causing the voltage at the terminal of the CPL (V_{CPL}) to rise over the rated value, as shown in Figure 5 V_{CPL} (t = 0.5s to 0.75s). The DVR is turned on again at 0.75s, and this time it absorbs the extra voltage (V_{DVR} = 100V) to regulate the voltage at the CPL's terminal to the rated value. Since the DVR is unable to regulate the voltage using solely reactive power, therefore, both real (d-axis component V_d) and reactive (q-axis component V_q) components are used. Also, the DVR is absorbing the excess power in this case, so it can store it in the battery and use it whenever it is needed. The waveforms in Figure 6 are used to investigate the effect of a load connected at the feeder's end being perturbed. With I_{RES} = 10A, the penetration of RES is kept constant throughout the simulation. From t = 0 to 0.5s, the load connected at the end of the feeder is approximately 6.25kVA (3.75kW CPL load and 2.5kVA linear constant impedance load). The RES generates power that is close to the demand from local loads, therefore there isn't much voltage drop across the CPL load. The DVR can control the marginal variation from the rated value with the q-axis component only (V_{DVR} = 15V) from 0.25 to 0.5s. This is further supported by the zoomed waveform from 0.15s to 0.4s. From 0.5 to 1s, the load connected rises significantly with the CPL load remaining the same at 3.75kW, while the constant impedance load rose to 7.5kVA. As the local demand increased beyond the capacity of RES generation, the variation from the rated voltage rose significantly, requiring the DVR to regulate the voltage using a combination of real and reactive power with V_{DVR} = 100V, as shown in Figure 6 from 0.75 to 1s. This is backed even more by the zoomed waveform. As the DVR supplies real power into the system, it is depleting its real power stock that it had built up during the excess RES generation. As a result, the voltage at the CPL terminal is regulated.

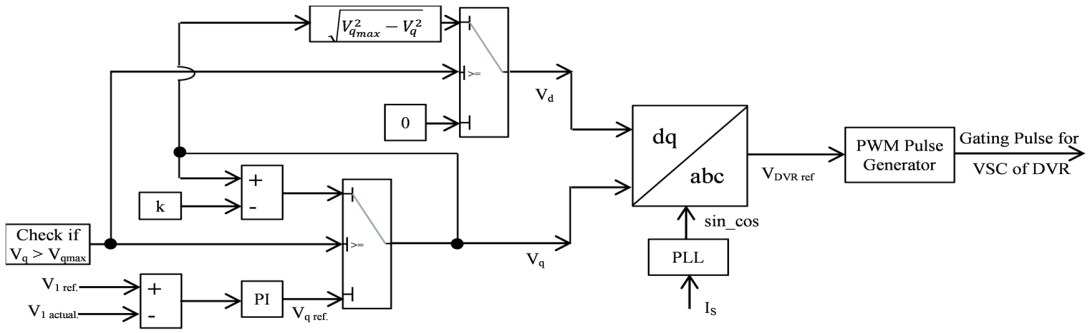


Fig. 3. Block diagram of the control scheme of the DVR.

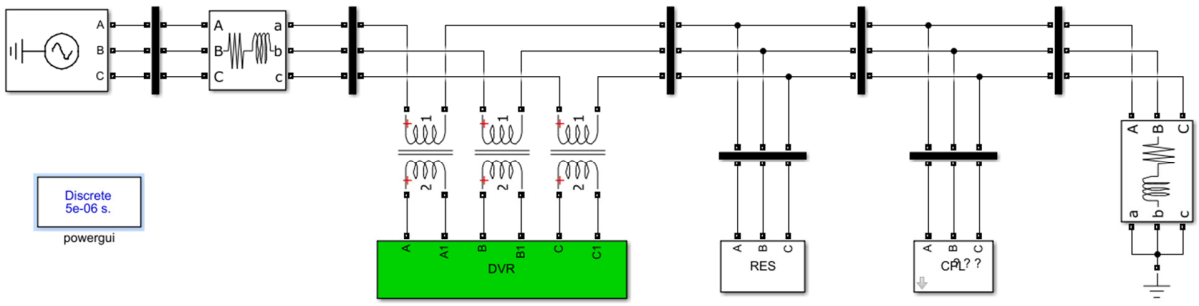


Fig. 4. MATLAB/Simulink simulation diagram of the system under consideration.

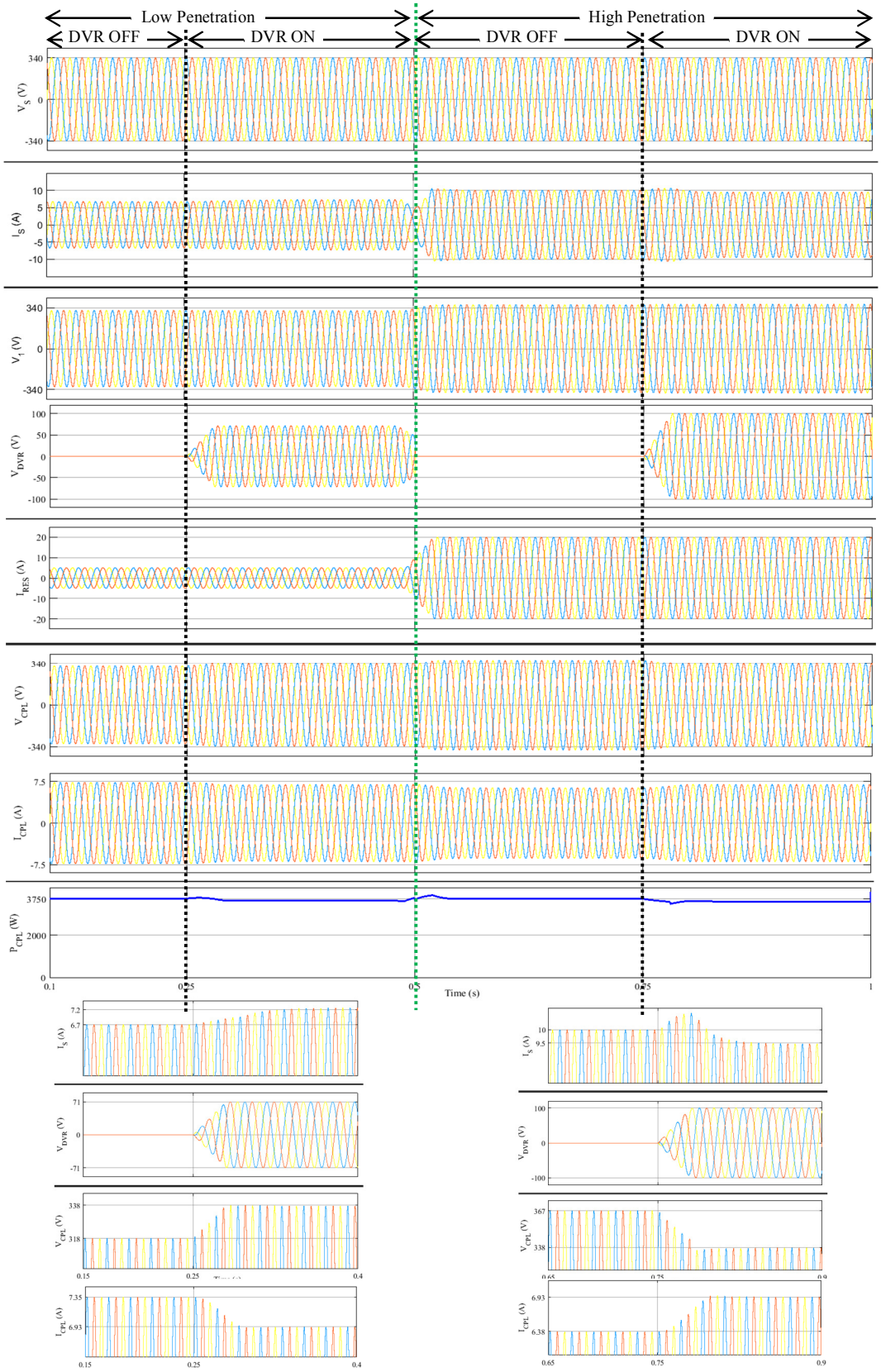


Fig. 5. Voltage and current waveform showing the efficacy of DVR during intermittency of the RES.

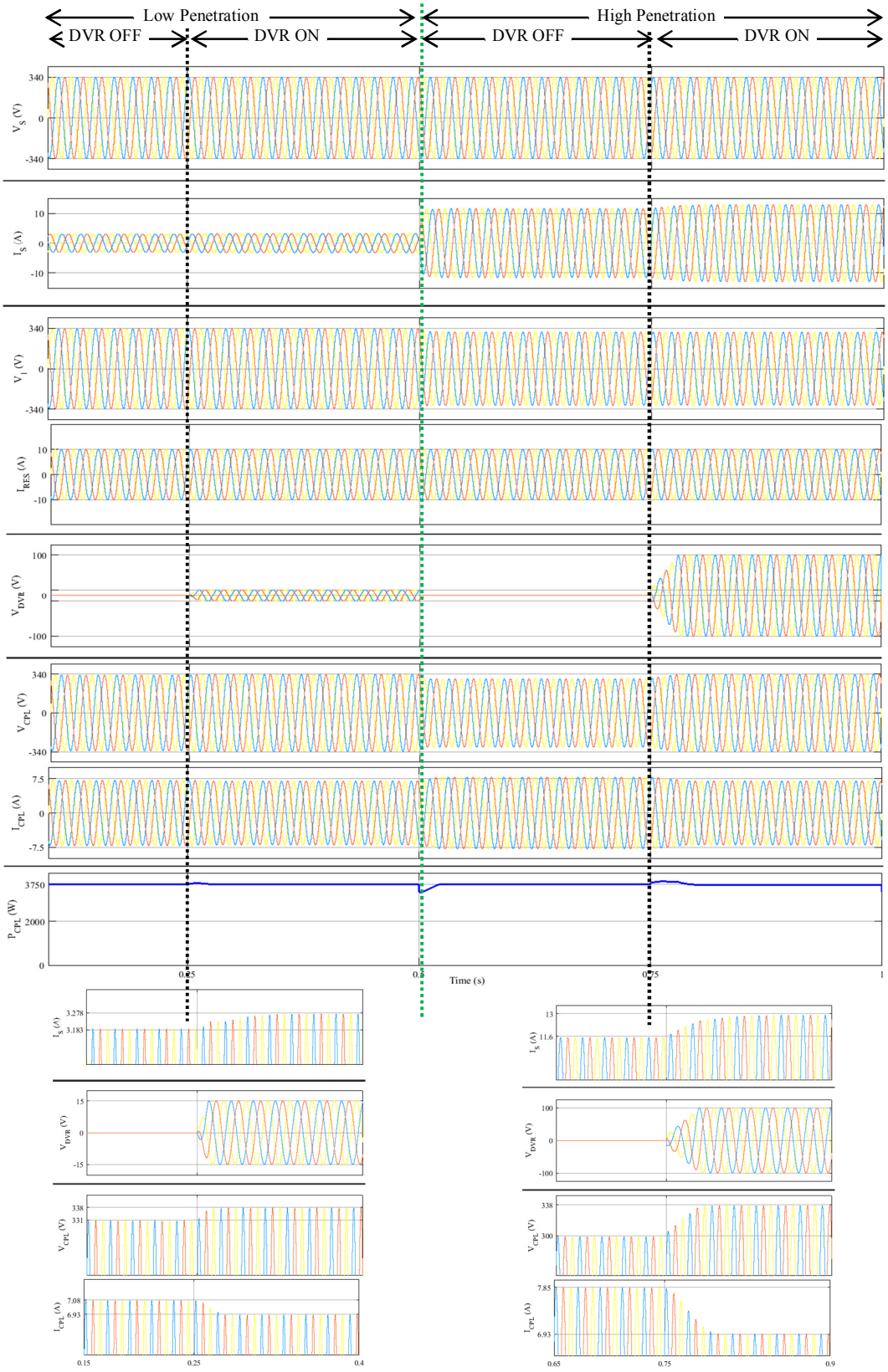


Fig. 6. Voltage and current waveform for showing the efficacy of DVR during load perturbations.

VI. CONCLUSION

In the presence of intermittent RES, voltage regulation in a radial microgrid feeding a sensitive CPL load has been investigated and demonstrated. Analysis was carried out to determine the value of the voltage at the CPL terminal in the absence of the DVR, as well as the value of voltage to be injected by the DVR. The Matlab/Simulink environment was utilized to simulate and control the proposed system. The simulation results for the intermittency of the RES source and the perturbation of load on the feeder were used to establish the performance of the DVR for voltage regulation. The demonstrated simulation results ensure that the voltage at the CPL terminal is relatively unaffected and settles in a short time. This allows the microgrid to benefit from greater RES penetration without compromising the operation of sensitive CPL loads.

REFERENCES

[1] D. Karimipour and F. R. Salmasi, "Stability Analysis of AC Microgrids With Constant Power Loads Based on Popov's Absolute Stability Criterion," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 7, pp. 696–700, Jul. 2015, https://doi.org/10.1109/TCSII.2015.2406353.

[2] D. P. Ariyasinghe and D. M. Vilathgamuwa, "Stability analysis of microgrids with constant power loads," in *2008 IEEE International Conference on Sustainable Energy Technologies*, Singapore, Aug. 2008, pp. 279–284, https://doi.org/10.1109/ICSET.2008.4747017.

[3] T. V. Krishna, M. K. Maharana, and C. K. Panigrahi, "Integrated Design and Control of Renewable Energy Sources for Energy Management," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5857–5863, Jun. 2020, https://doi.org/10.48084/etasr.3613.

[4] V. Verma and R. Gour, "OLTC-DVR hybrid for voltage regulation and averting reverse power flow in the micro-grid with intermittent renewable energy sources," in *2016 IEEE Industrial Electronics and Applications Conference (IEACon)*, Kota Kinabalu, Malaysia, Aug. 2016, pp. 81–87, https://doi.org/10.1109/IEACON.2016.8067360.

[5] A. Emadi, A. Khaligh, C. H. Rivetta, and G. A. Williamson, "Constant power loads and negative impedance instability in automotive systems: definition, modeling, stability, and control of power electronic converters and motor drives," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 4, pp. 1112–1125, Jul. 2006, https://doi.org/10.1109/TVT.2006.877483.

[6] J. M. Coulomb, "Stability analysis and optimal control design for AC-DC power system with constant power load," University of Pittsburgh, Pittsburgh, PA, USA, 2013.

[7] J. von Appen, M. Braun, T. Stetz, K. Diwold, and D. Geibel, "Time in the Sun: The Challenge of High PV Penetration in the German Electric Grid," *IEEE Power and Energy Magazine*, vol. 11, no. 2, pp. 55–64, Mar. 2013, https://doi.org/10.1109/MPE.2012.2234407.

[8] S. G. Abhyankar and A. J. Flueck, "Simulating voltage collapse dynamics for power systems with constant power load models," in *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, Pittsburgh, PA, USA, Jul. 2008, pp. 1–6, https://doi.org/10.1109/PES.2008.4596587.

[9] V. Verma and R. Gour, "Step-less voltage regulation on radial feeder with OLTC transformer-DVR hybrid," in *2015 6th International Conference on Power Electronics Systems and Applications (PESA)*, Hong Kong, China, Sep. 2015, pp. 1–6, https://doi.org/10.1109/PESA.2015.7398912.

[10] A. Ghosh and G. Ledwich, "Compensation of distribution system voltage using DVR," *IEEE Transactions on Power Delivery*, vol. 17, no. 4, pp. 1030–1036, Jul. 2002, https://doi.org/10.1109/TPWRD.2002.803839.

[11] S. Suraya, S. M. Irshad, M. F. Azeem, S. F. Al-Gahtani, and M. H. Mahammad, "Multiple Voltage Disturbance Compensation in Distribution Systems using DVR," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5732–5741, Jun. 2020, https://doi.org/10.48084/etasr.3485.

[12] A. H. Soomro, A. S. Larik, M. A. Mahar, A. A. Sahito, and I. A. Sohu, "Simulation-based Analysis of a Dynamic Voltage Restorer under Different Voltage Sags with the Utilization of a PI Controller," *Engineering, Technology & Applied Science Research*, vol. 10, no. 4, pp. 5889–5895, Aug. 2020, https://doi.org/10.48084/etasr.3524.

[13] P. Jayaprakash, B. Singh, D. P. Kothari, A. Chandra, and Kamal-Al-Haddad, "Control of Reduced Rating Dynamic Voltage Restorer with Battery Energy Storage System," in *2008 Joint International Conference on Power System Technology and IEEE Power India Conference*, New Delhi, India, Jul. 2008, pp. 1–8, https://doi.org/10.1109/ICPST.2008.4745199.

[14] A. Engler and N. Soultanis, "Droop control in LV-grids," in *2005 International Conference on Future Power Systems*, Amsterdam, Netherlands, Aug. 2005, https://doi.org/10.1109/FPS.2005.204224.

J. Inst. Eng. India Ser. A
https://doi.org/10.1007/s40030-022-00675-9

ORIGINAL CONTRIBUTION

Wind-Induced Interference Effects on Low-Rise Pitched Roof Structure Based on CFD Simulations

Amartya Sinha[1] · Aditya Kumar Jha[1] · Ritu Raj[1]

Received: 16 December 2021 / Accepted: 23 July 2022
© The Institution of Engineers (India) 2022

Abstract This work focuses on the evaluation of interference factors and the study of the variation of wind effects on low-rise pitched roof structures with one side kept open to the wind environment under various interference conditions induced by changing the spacing in full blockage configurations. Since no provisions are available in the wind standards for use in designing such structures, the numerical simulations are performed using Ansys CFX fluid flow solver through a standard $k - \varepsilon$ turbulence model. Such type of study requires the use of either wind tunnel pressure measurements, or the numerical simulation of the wind environment virtually on a computational fluid dynamics (CFD) domain. In this study, the CFD simulations have been carried out for four different spacing conditions and seven different wind incidence angles varied at the interval of 30° to obtain the wind pressure coefficients (C_p). Using the obtained results, the interference difference (ID) and the interference factor (IF) for each of the considered configurations were determined. The variations of IF along the roof direction were also analyzed. Also, the drag and lift force coefficients were calculated, and their variations were studied. It was found that the increase in spacing and the change in wind incidence angles affected the interference parameters. The wind flow pattern around and inside the structure varied for different wind incidence angles. The formation of eddies at various stagnation regions affected the drag and lift forces acting on the structure.

Keywords Pitched roof · Wind interference · Shielding effect · Ansys CFX · Interference factor · Interference difference

Introduction

Construction of building structures is widely required all over the world as a result of technological breakthroughs and the rising demand for infrastructure. Buildings are subjected to variable loads on their side walls and roofs due to the presence of wind flow.

In building structures, the roof is a very critical functional component of the building. Generally, in low-rise structures, wind load exerts extreme outward pressure on the roof, which in the case of high-speed wind flow, often results in blowing off of the roof. Both the roof form and the roof slope are essential factors in the structure's wind load resistance. Apart from these, the arrangement of the nearby buildings in terms of blockage offered by them and the formation of the zones of separation and reattachment of the wind flow can also change the wind load acting on the structure. This phenomenon is referred to as the interference effect of the wind and is commonly seen in urban locations where clusters of structures are constructed at a location. When multiple structures are under the consideration, the complexity of the study increases due to a large number of datasets available. Therefore, the interference factor, which is the ratio of the coefficient of pressure at a particular structure configuration to the coefficient of pressure on a similar building under isolated conditions, is a valuable tool for studying the change of the wind action and the pressure on the structure.

The effects of wind load acting on a structure can be studied either with the help of tests on a model in a boundary

✉ Ritu Raj
rituraj@dtu.ac.in

[1] Department of Civil Engineering, Delhi Technological University, Delhi, India

Published online: 16 August 2022

layer wind tunnel or with the help of a virtual wind tunnel created on a computational fluid dynamics (CFD)-based software such as Ansys CFX fluid flow package in which a 3D model of the structure placed in a wind domain is subjected to a natural wind flow profile and properties like flow velocity, pressures, density and temperature at various locations are determined. This results in a cost-effective and quick analysis of the flow.

The effects of wind load are very variable and depend upon the geometric properties of the structure as well as the wind profile and terrain. On various types of structures, various analytical and numerical researches have been conducted. Wind flow measurements were carried out on actual low-rise buildings of various shapes and side ratios by Blackmore (1988), Gerhardt & Kramer (1992), Meecham (1992) and Mehta et al. (1992), who discovered that small changes in the edges and shapes of the walls have a significant impact on the wind flow [1–4].

Wind load studies on the high rise as well as on low-rise structures of different heights and geometries have been carried out by many researchers where it was found that different plan shapes of the building can lead to completely different wind actions. Amin and Ahuja (2011), Paul and Dalui (2016), Mallick et al., (2019), Bhattacharyya et al. (2014), Chakraborty et al. (2014), Mukherjee et al. (2014), Bairagi and Mukherjee (2017), Sanyal and Dalui (2021), Gaur and Raj (2021) and Kumar and Raj (2021) have worked on high rise buildings of various shapes like L, T, Z, C, E, +, N, Y, Octagonal–Oval, etc., by using experimental as well as numerical studies through which wind pressure patterns were studied [5–15].

CFD simulations are also widely used by researchers in studies for wind load effects on other roof and plan shapes. The effect of wind direction and roof slope for single and double-storied pyramidal roof buildings has been studied by Singh and Roy (2019 & 2021) [16–19]. Dome-shaped structures and single and double-span cylindrical roofs have also been studied by Zhou, et al. (2019) and Verma and Ahuja (2015) [20–22].

In the case of clustered building structures, the building constructed on the leeward side may experience shielding from the wind by the windward building or may also experience enhanced wind load, depending upon the configuration of the structures and their geometries and the modified wind flow around the structures caused by the presence of interfering structure. The wind tunnel tests on hip roofs were carried out by Ahmad and Kumar (2001) where it was observed that interference effects offer significant enhancements and shielding effects at different configurations [23]. The interference effect on corner configured structures using CFD for different geometry and blockage configurations was studied by Gaur et al. (2021) [24]. Experimental studies on wind-generated proximity effects were carried out by Nagar et al. (2020 and 2021) between 2 plus-shaped, 2 square and 2 H-shaped tall buildings in different blockage configurations [25, 26]. Square and fish plan-shaped structures have also been experimentally studied for the action of wind interference effect in the boundary layer wind tunnel by Pal and Raj (2021) by employing two parameters which are the interference factor and the interference difference [27]. The buffeting action of wind and its effect on the low rise structures have been studied by Stathopoulos (1984) [28]. Interference effects on chimneys have also been studied by Shruti et al. (2019) by using the interference factor as the parameter [29].

The set of Indian Standards requirements on the estimate of wind loads on structures, IS 875 (Part 3): 2015, was thoroughly reviewed, and the provisions were referred to at various stages [30].

In this paper, the effect of change in spacings of the buildings in full blockage condition on the wind pressure acting on the roof and the drag and lift forces acting on a low-rise rectangular pitched roof building with one side kept open is studied by carrying out CFD simulations.

The effect on the building in the vicinity of another building can be of two natures, either the wind load on the building will enhance or will get reduced. The former is known as the wind load enhancement effect, and the latter is known as the shielding effect. The effect which will dominate depends upon the relative placement and orientation of the structures with the wind flow direction. This CFD study helps to quantify the wind-induced interference effects under different spacing configurations and wind flow directions.

Structure Under Study

For the present study, a low-rise rectangular plan pitched roof structure model is considered as the instrumented model on a 1:10 geometric scale. The pitched roof angle is 20° with horizontal and maximum and the minimum wall lengths are taken as 1 and 0.6 m, respectively. The wall height of the structure is 0.5 m. A full wall opening has been provided in one of the smaller walls in the structure as shown in Fig. 1.

Equation (1) represents the profile of the pitched roof:

$$z = 0.6092 - 0.364|x| \tag{1}$$

The wall thickness of the models is taken as 0.01 m. Since the structure is of small height, the models are subjected to a uniform wind velocity of 10 m/s as the wind flow around the structure lies in the atmospheric surface layer zone, in which, for simplicity in the analysis, a uniform velocity profile can be assumed.

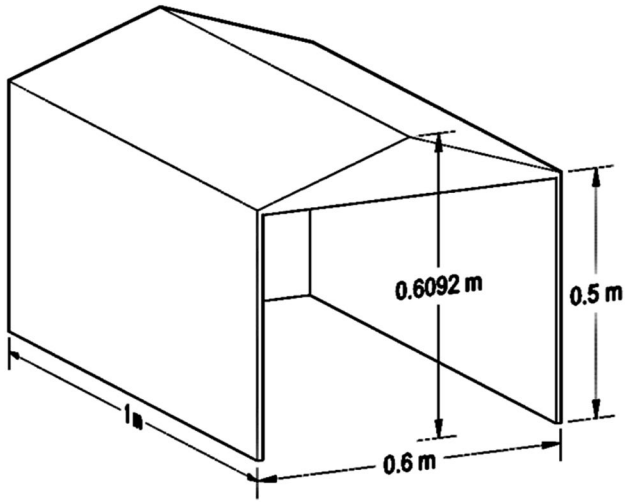


Fig. 1 Model of the structure

Numerical Simulation Methodology

Various commercial software available through which CFD simulations can be performed. Ansys [31] is one of the most widely used software suites comprising multiple engineering simulation software tools. For CFD simulations for wind flow around bluff bodies, the simplest solver is the Ansys CFX solver. Therefore, the simulations for the present study are performed in the Ansys 2021 R2 CFX fluid flow package at high turbulence intensity in the $k-\varepsilon$ turbulence model. This model is the most common turbulence model used for CFD studies on wind flow around bluff bodies. Here, k denotes the turbulent kinetic energy, and ε denotes the rate of dissipation of the turbulent kinetic energy in the form of eddies generated in the wind due to the presence of a bluff body.

The CFD simulations are done using five basic steps, which are:

- Model geometry creation,
- Flow domain setup,
- Domain discretization or meshing,
- Setup of flow physics and boundary conditions,
- Generation of solutions

Modelling of the structure under study was done using a 3D modelling software, in the present case, AutoCAD. Thereafter, the creation of a fluid flow domain around the structure is required, which will act as a virtual wind tunnel. This was done through the DesignModeler tool provided by Ansys.

In CFD simulations, the flow domain is discretized into several smaller elements, also known as the meshing of the flow domain. These elements can be of different shapes like tetrahedron, hexahedron, prism, or pyramidal. The continuity of the flow and the momentum of the flow are balanced along all the nodes of the discretized domain based on the Reynold's averaged Navier–Stokes equation. In the present study, tetrahedral meshing was adopted. In the meshing process, the sizes of the elements need to be adjusted to obtain a good quality mesh.

The next step involves the boundary conditions setup and specification of the wind profile which will exert the load on the structure. The side walls and the top wall of the domain were defined as free slip walls ($\tau_{\text{wall}}=0$ & $V_{\text{wall}}\neq 0$), where τ_{wall} is the shear stress at the wall, and V_{wall} is the velocity along the wall surface. But, the component of velocity normal to the wall surface is 0. This is done to replicate natural freestream flow conditions at the boundaries of the virtual wind tunnel. The model faces and the ground were defined as no-slip walls ($V_{\text{wall}}=0$) which means the component of fluid velocity along the boundary of the wall is zero. The no-slip wall condition ensures that wind flow is subjected to viscous resistance from the ground surface and the faces of the structure. Relative pressure at the outlet was defined with 0 Pa. In the present study, uniform homogeneous wind flow with a wind velocity of 10 m/s is provided at the inlet of the domain.

After the flow setup is complete, the Ansys solution manager performs multiple iterations of solving the continuity and momentum equations at each node of the mesh. As the iterations proceed, the residual values of mass and momentum along each direction reduce and the solution converges to the final stage.

Once the solution is complete, the results can be accessed through the CFD post-tool of Ansys. Through the result file, the complete flow around the body can be visualized with the help of streamlines and the pressure and forces acting over the faces of the body can be found. The magnitudes of the pressure and forces will depend upon the wind velocity as well as the model size. To represent the pressures and forces in general terms, a dimensionless quantity is required. The pressure coefficient is a dimensionless quantity defining the relative pressure at any point w.r.t the design pressure. Mathematically, it is defined as with the help of Eq. (2) [32].

$$\text{Pressure Coefficient} = C_{\text{p}} = \frac{P - P_o}{\frac{1}{2}\cdot\rho\cdot V^2} \tag{2}$$

where P is the wind pressure acting at any point, P_{o} is the reference pressure which is the pressure in the region outside the influence of the building, i.e. outlet, defined as 0 Pa in the present study, and ρ & V are the density of air and design wind speed, respectively.

The expression for C_{p} can be created in the inbuilt function calculator of Ansys using Eq. (2) to express the wind pressure acting on the faces of the structure in dimensionless

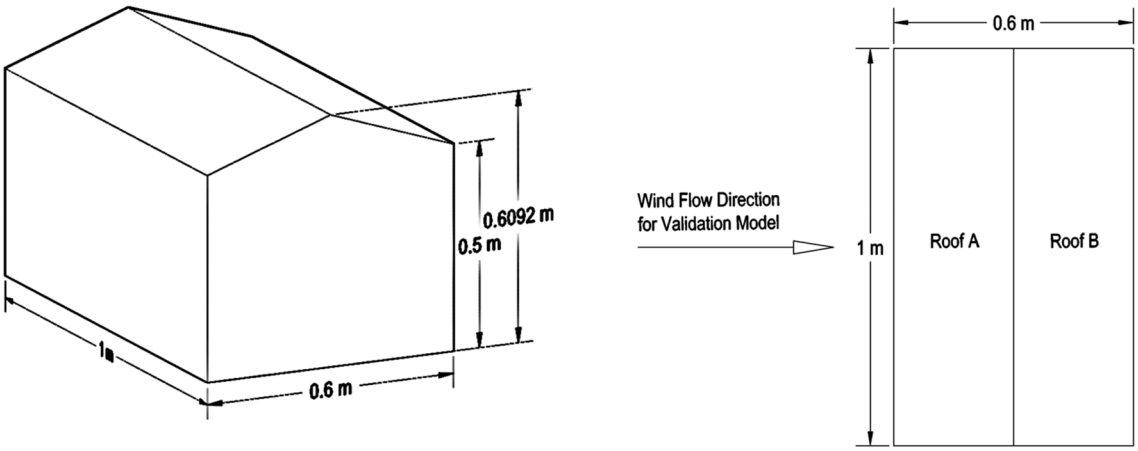


Fig. 2 Validation model

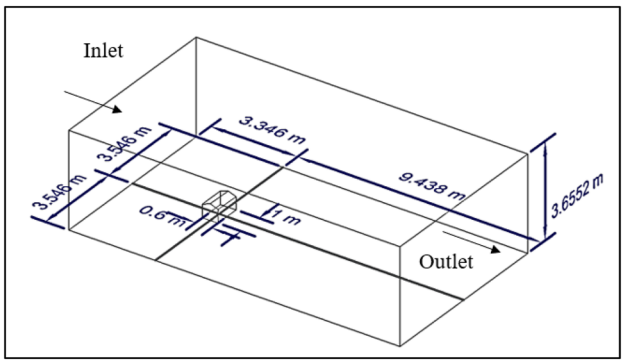


Fig. 3 Wind flow domain

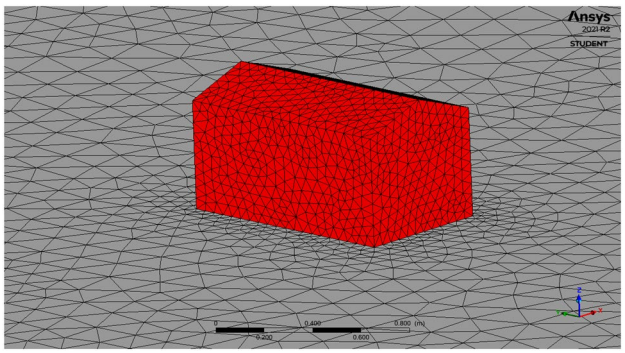


Fig. 4 Mesh diagram

terms. The values of C_p can also be plotted over the whole face in the form of C_p contours to visualize the variation of C_p over the different faces.

Validation with Indian Standards

A validation model with the same external dimensions and meshing as the flow setup was subjected to the same normal wind velocity of 10 m/s at the inlet to validate the flow setup as shown in Fig. 2. The validation of the setup was done by comparing the coefficient of pressure (C_p) values for closed pitched roof structure obtained from Ansys CFX to those obtained from the IS code (based on wind tunnel tests).

The recommendations on CFD applications in wind engineering by Franke et al. (2004) were followed for determining the domain size. Figure 3 depicts the dimensions adopted for the wind flow domain. The inlet and outlet faces were kept at a clear distance of 5H and 15H, respectively, from the building. The side and top clearance from the building were also taken as 5H. Here, H is the height of the structure and is equal to 0.6092 m.

The volume mesh size is defined at 0.2 m, while the mesh size at the building sides has been gradually lowered to 0.05 m, the diagram for which is shown in Fig. 4.

The mesh quality in a CFD simulation can be judged on various mesh metrics like skewness, orthogonal quality, aspect ratio, smoothness, etc. A bad mesh quality can cause convergence difficulties, resulting in a larger error accumulation. A good quality mesh results in capturing fine geometric details of the flow and therefore, a more accurate solution. In a good quality mesh, the elements should have similar shape and size and any reduction or expansion in size should be gradually applied. Ansys recommends orthogonal quality and skewness as the most important metric for judging mesh quality [33]. Table 1 shows the recommended mesh metric values from Ansys for skewness and orthogonal quality. A low value of orthogonal quality and a high value of skewness result in a poor quality of the mesh. In 3D, most cells should be good in these metrics; however, a small percentage of the cells can also fall in the poor range.

Table 1 Recommended values for skewness and orthogonality controls in meshing

Skewness					
Excellent	Very good	Good	Acceptable	Bad	Unacceptable
0–25	0.25–0.50	0.50–0.80	0.80–094	0.95–0.97	0.98–1.00
Orthogonal quality					
Unacceptable	Bad	Acceptable	Good	Very good	Excellent
0–0.001	0.001–0.14	0.15–0.20	0.20–0.69	0.70–0.95	0.95–1.00

Table 2 Validation of CFD Setup from IS 875 (Part 3): 2015

Wind angle (θ°)	C_{pe} for surface from CFD simulation		Average value of C_{pe} for roof from CFD simulation	Average C_{pe} for Roof as per IS 875 (Part 3)	Average % Error
	Roof A	Roof B			
0	−0.868	−0.530	−0.699	−0.6	16.5%

The minimum, average and maximum values of skewness are 0.00064, 0.20 and 0.84, respectively. The minimum, average and maximum values of orthogonal quality are 0.15, 0.79 and 0.99, respectively.

Therefore, in the adopted mesh, the average skewness was in the excellent range and the orthogonal quality was in the very good range. Also, no element was under bad and unacceptable ranges of skewness and orthogonal quality. Therefore, the mesh was used for the simulations in this study.

Uniform wind flow was simulated along the direction perpendicular to the longitudinal axis of the model which corresponds to the 0° wind incidence in IS 875 (Part 3): 2015 and for the whole roof, an average percentage error of 16.5% was measured shown in Table 2, on comparing the result obtained for the validation model with the values given in Table 6 of the IS 875.

The wind escaping from the sharp edges of the bluff body results in the development of suction in the regions close to the edges. Since, the CFD simulations capture the smallest variations of the wind pressure comparable to the size of the elements of the meshing, the high suction values near the edges of the building slightly increase the average suction C_{pe}. In the present study, from the comparison with IS 875, a difference of 0.099 is observed in the values of C_{pe}.

Since the results obtained on the validation model are within satisfactory limits of the codal provisions, the same meshing and boundary setup have been adopted for the simulations carried out in this study.

The simulations are carried out for different spacings between the instrumented model and the interfering model and different wind incidence angles ranging from 0° to 180° at an interval of 30° as shown in Fig. 5. The interfering model in the present study is taken as the model having the same external dimensions as the instrumented model. The instrumented model is kept exactly behind the interfering model, i.e. in full blockage condition, in 0° wind incidence and the spacings between the two are varied in terms of the width of the building, b (= 0.6 m). The four different spacing configurations (S) considered in this study are corresponding to $\frac{b}{2}$, b, $\frac{3b}{2}$ and $2b$, values for which are indicated in Fig. 6a to d.

Average Pressure Coefficients

These graphs in Fig. 7 represent the $C_{p\,mean}$ variation in pitch roof faces A & B w.r.t wind incidence angles. For the external roof face, the $C_{p\,mean}$ graph has a similar variation and is suction in nature for each of the spacings. The isolated structure case also shows a similar variation; however, the suction pressure, particularly at 0°, shows a higher value of about −0.4 for both the faces A & B,

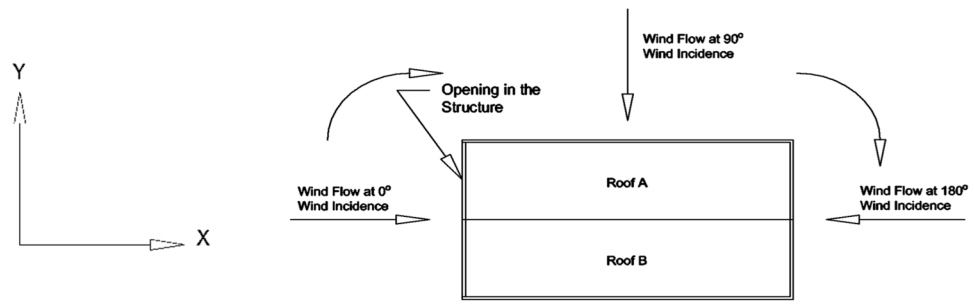


Fig. 5 Variation of wind incidence angles

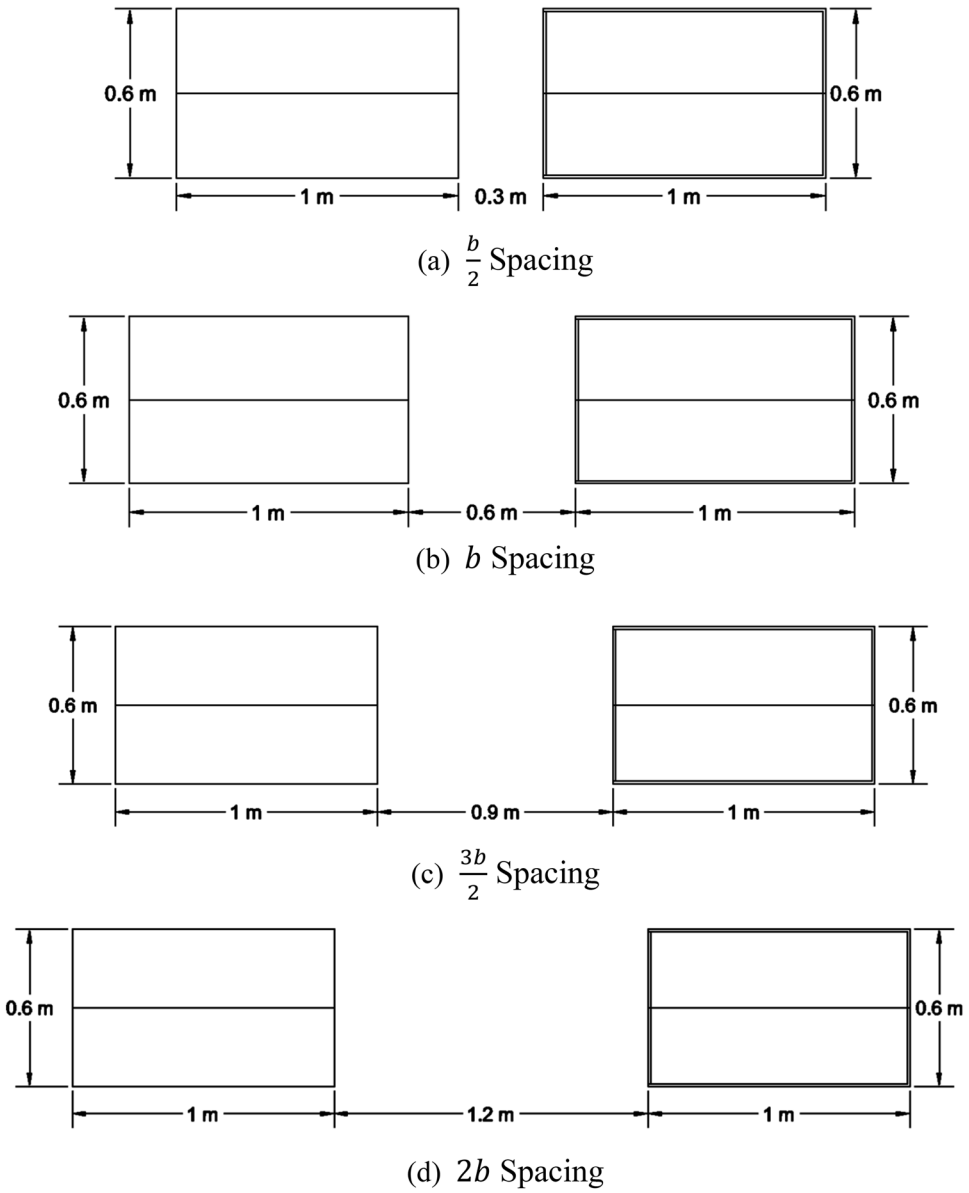


Fig. 6 Spacing configurations

whereas the presence of the interfering building in the other cases results in eddy formations which leads to a reduction in the suction pressure.

Similarly, the $C_{\text{p mean}}$ curves for different spacings for the internal face of both faces A & B show similar variation with a gradual increase in the $C_{\text{p mean}}$ values for each spacing. This is due to the free wind flow action on the internal face of the roof with the increase in spacing. The isolated structure condition understandably shows a higher $C_{\text{p mean}}$ value of about 1 for 0° because of the free wind incidence on the roof from the open side of the buildings. The presence of the interfering building causes a reduction in $C_{\text{p mean}}$ values for the interference cases due to its shielding effect. The $C_{\text{p mean}}$ values gradually fall with each wind incidence angle as the direct entry of wind streamline is restricted with each increment in wind incidence angle. As a result, negative $C_{\text{p mean}}$ values are eminent post 90° wind incidence.

Interference Parameters

A variety of research has been executed in the area of wind-induced interference effects. As per Khanduri et al. (1998), the interference effects require systematic studies since the available studies cannot be compared directly due to the different geometrical configurations [34]. Therefore, in the present study, the change in the wind action with spacing and its effects on the interfering buildings are studied through two interference parameters which are the interference factor (IF) and the interference difference (ID). IF is defined as the

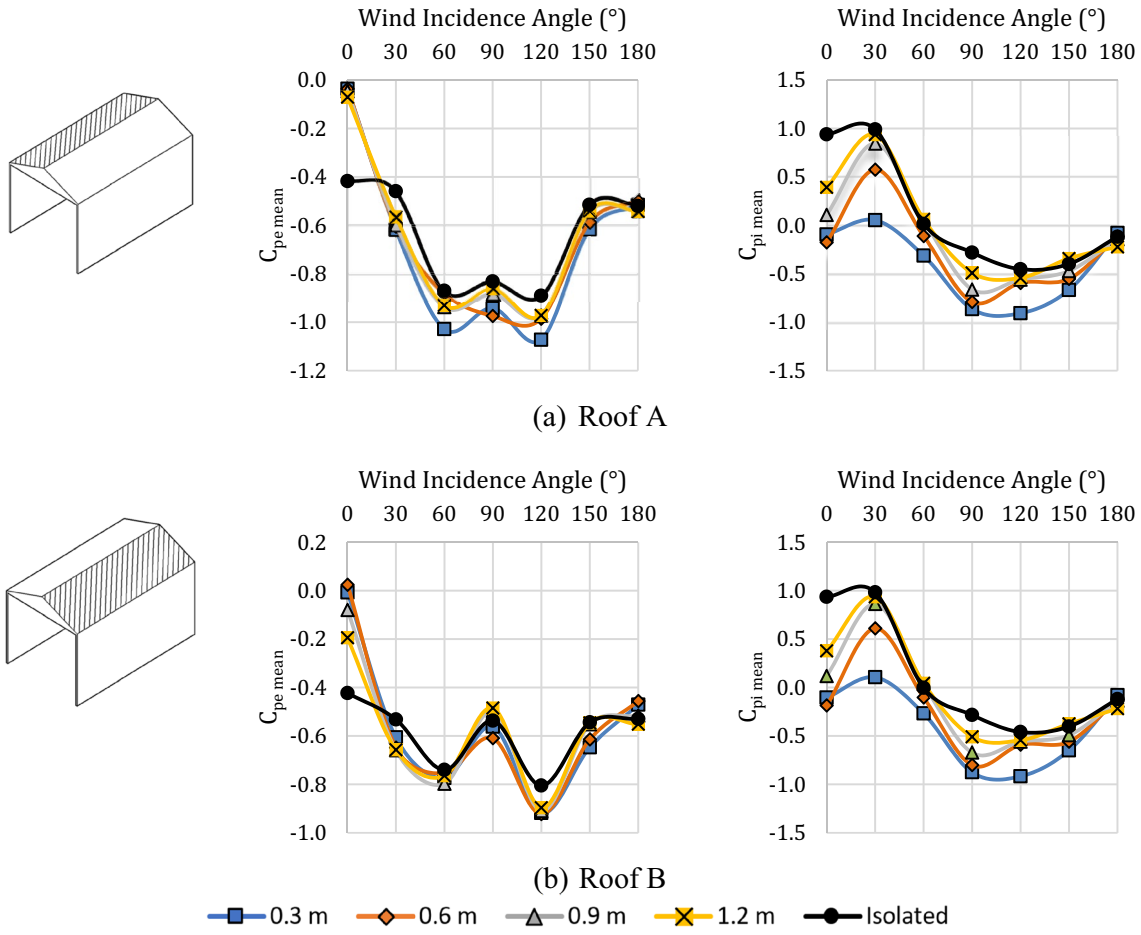


Fig. 7 Variation of average pressure coefficient with wind incidence angles

ratio of the C_p for a building with the interfering building to the C_p for the same building under isolated conditions. ID is defined as the difference between the absolute value of C_p for a building with the presence of the interfering building and the absolute value of C_p for the same building under isolated conditions. These parameters help to numerically represent the effect presented by the interfering building to the wind action on the building under study.

The values of IF may be positive as well as negative. Positive IF implies that the nature of the wind action, i.e. pressure, or suction is the same in the isolated and the interference cases, whereas negative IF implies that the nature of the wind action has changed. If the magnitude of the IF is less than 1, it can be suggested that the interfering building has shielded the building under the study from the wind action. On the other hand, if the magnitude of IF is more than 1, enhancement of the wind load is said to be taking place due to the presence of the interfering building. However, since IF is a ratio, the factor can be misleading if the values of coefficients of pressure are in extreme ranges. Therefore, to effectively study the interference effects, the use of ID is also necessary. ID can represent the change in the C_p more effectively even when extreme values of C_p are observed. The IF gives significant results when the interference effects of localized regions are to be analyzed.

Therefore, in the present study, the IDs in different spacing configurations and different wind incidence angles have been calculated w.r.t mean pressure coefficients on external and internal faces of the roofs separately which are presented in Tables 3 and 4. The variation of IF along the roof length is studied by considering 9 points along the longitudinal direction of the structure along the centre of the roof surfaces (labelled 1 to 9) for each of the Roof A and Roof B as shown in Fig. 8 for different spacing configurations and wind incidence angles. The spacing between the two consecutive points is equal to 0.1 m. The values of IFs have been calculated by considering the net coefficient of pressure, given by $C_{pe} - C_{pi}$ at each of these 9 points on the roof.

The negative values of ID in Tables 3 and 4 signify the shielding effect presented by the interfering building. The 0° wind incidence, where the structure is completely shielded, has significant negative IDs and it can be observed that the

Table 3 Interference difference for pitched roof surface A

Wind incidence angle	External surface				Internal surface			
	0.3 m	0.6 m	0.9 m	1.2 m	0.3 m	0.6 m	0.9 m	1.2 m
0°	−0.38	−0.38	−0.37	−0.35	−0.85	−0.77	−0.83	−0.55
30°	0.16	0.14	0.14	0.11	−0.94	−0.42	−0.15	−0.06
60°	0.16	0.01	0.07	0.06	0.30	0.09	0.04	0.05
90°	0.11	0.14	0.05	0.03	0.58	0.51	0.38	0.21
120°	0.18	0.09	0.08	0.08	0.45	0.14	0.11	0.09
150°	0.10	0.07	0.03	0.02	0.27	0.15	0.07	−0.06
180°	0.00	−0.02	0.02	0.02	−0.04	0.00	0.02	0.10

Table 4 Interference difference for pitched roof surface B

Wind incidence angle	External				Internal			
	0.3 m	0.6 m	0.9 m	1.2 m	0.3 m	0.6 m	0.9 m	1.2 m
0°	−0.42	−0.40	−0.34	−0.23	−0.83	−0.75	−0.82	−0.56
30°	0.07	0.12	0.13	0.13	−0.88	−0.37	−0.12	−0.05
60°	0.03	0.01	0.06	0.02	0.27	0.10	0.03	0.04
90°	0.03	0.07	−0.04	−0.05	0.59	0.52	0.39	0.23
120°	0.11	0.12	0.10	0.09	0.46	0.13	0.10	0.08
150°	0.10	0.07	0.01	0.00	0.25	0.16	0.09	−0.03
180°	−0.06	−0.07	−0.01	0.02	−0.04	0.00	0.02	0.10

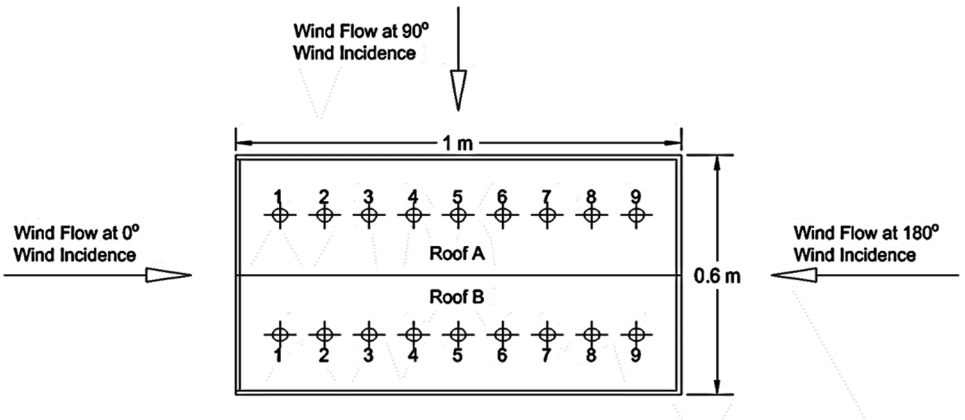


Fig. 8 Points along the longitudinal direction of the roof

IDs for external faces show a decreasing magnitude as the spacing is increased. As the spacing increases, the interference effect reduces and the flow around the instrumented building becomes more and more similar to the isolated configuration. The IDs for external faces at 90° wind incidence are very close to 0 since the wind strikes both instrumented model and the interfering model simultaneously and no significant shielding or enhancement is seen. The slight variation is due to the changed wind flow pattern near the building environment due to the introduction of the interfering model.

It can also be seen that the IDs for external surfaces are significantly smaller, barring the 0° wind incidence case due to complete blockage to the wind flow. On the other hand, a comparatively higher difference is seen in the IDs for internal surfaces, which have both positive as well as negative significant values for different wind incidences. The 0° and 30° configurations show a significant shielding effect which reduces, especially in the 30° configuration, as the spacing is increased. The other wind incidences have significant enhancement of wind load on internal surfaces, which sharply reduces, as the spacing is increased.

Figure 9 shows the variation of the interference factors along the length of the roof, based on the net pressure coefficient values. Since the net C_p values for the isolated structure are negative, the wind action, in the isolated condition, exerts net uplift pressure on the roof. For 0° configurations, similar variation is seen for Roof A and B as both

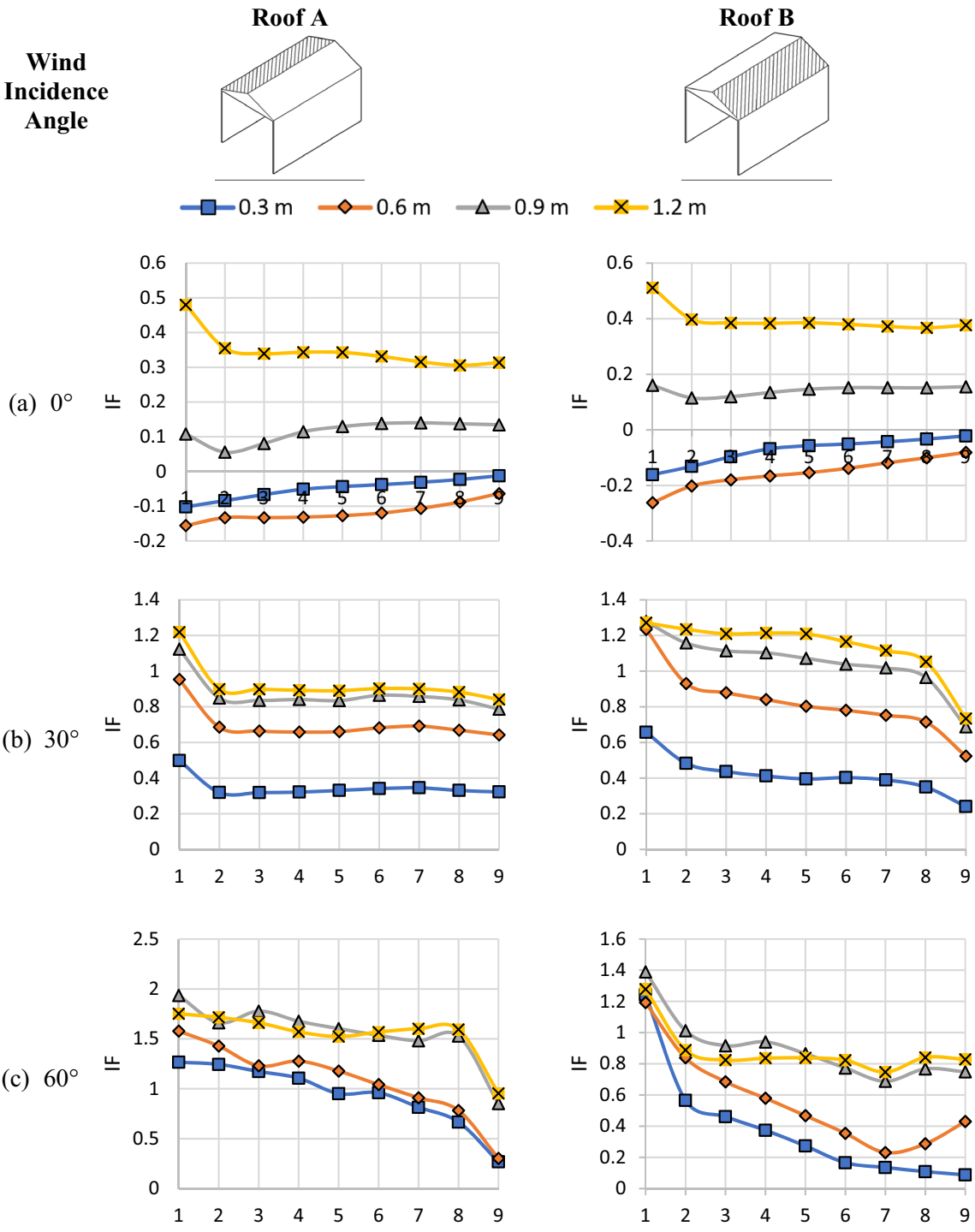


Fig. 9 Variation of interference factors along the roof

are symmetric to the wind flow and it is observed that the nature of the wind action changes in the 0.3 and 0.6 m spacing configurations since the IF is negative at all points along the roof. However, the nature of the wind action for 0.9 m and 1.2 m spacing configurations remains the same. As the spacing is very less in 0.3 and 0.6 m configurations, a proper wind flow is not developed in the internal environment of the building, as seen from the velocity contours in Fig. 10. As the spacing increases, internal uplift pressure increases and higher uplift pressure is observed for 0.9 and 1.2 m spacings. As the spacing is increased, the values of IF move towards unity and the interference effect reduces.

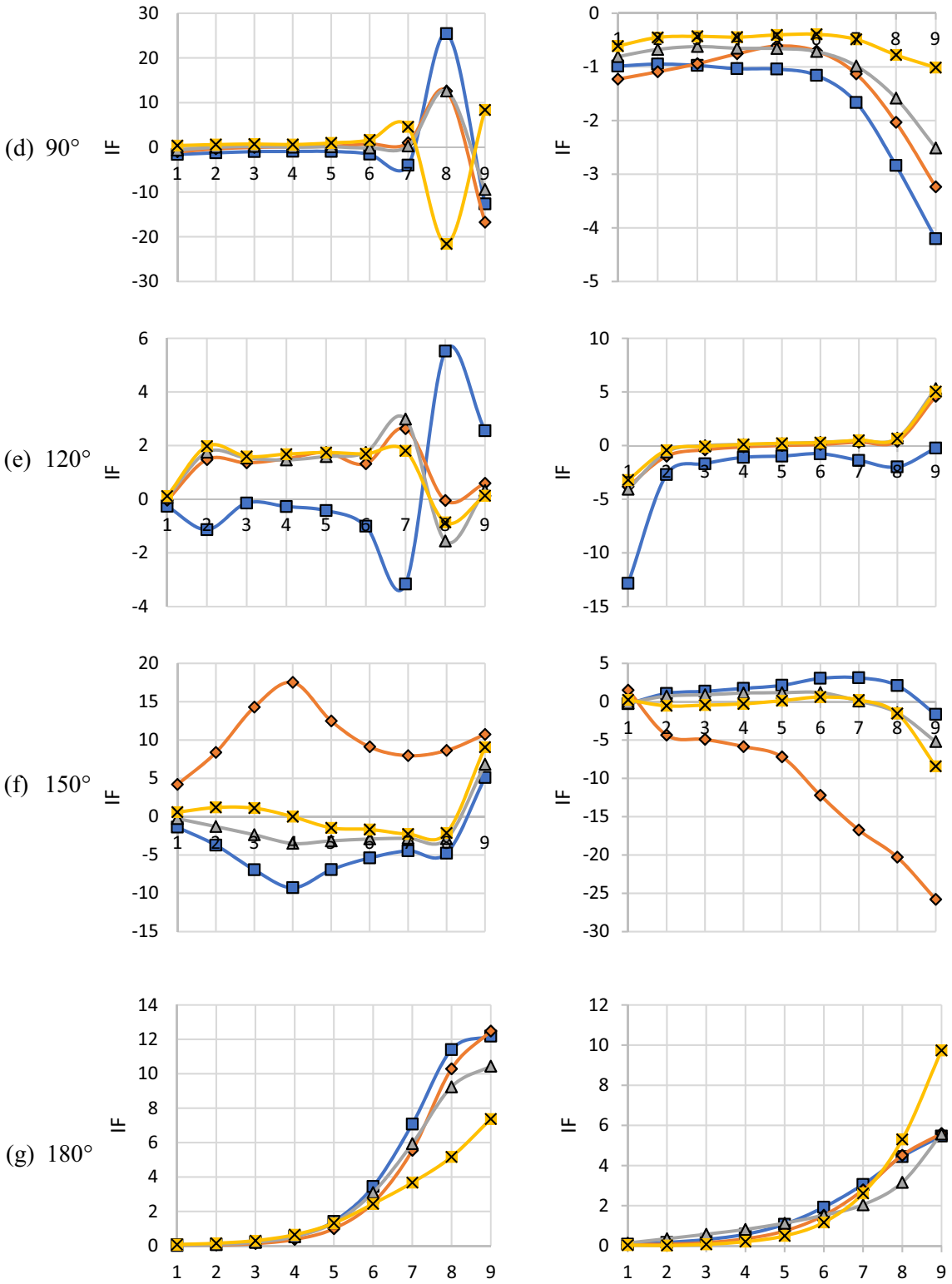


Fig. 9 (continued)

The four different spacings in the study correspond to different flow behaviours of wind around the structures. At the inlet, a uniform and homogeneous wind profile is provided with a velocity of 10 m/s. As the wind escapes from the roof of the interfering building, the flow accelerates in the neighbouring region which can be observed as a dark

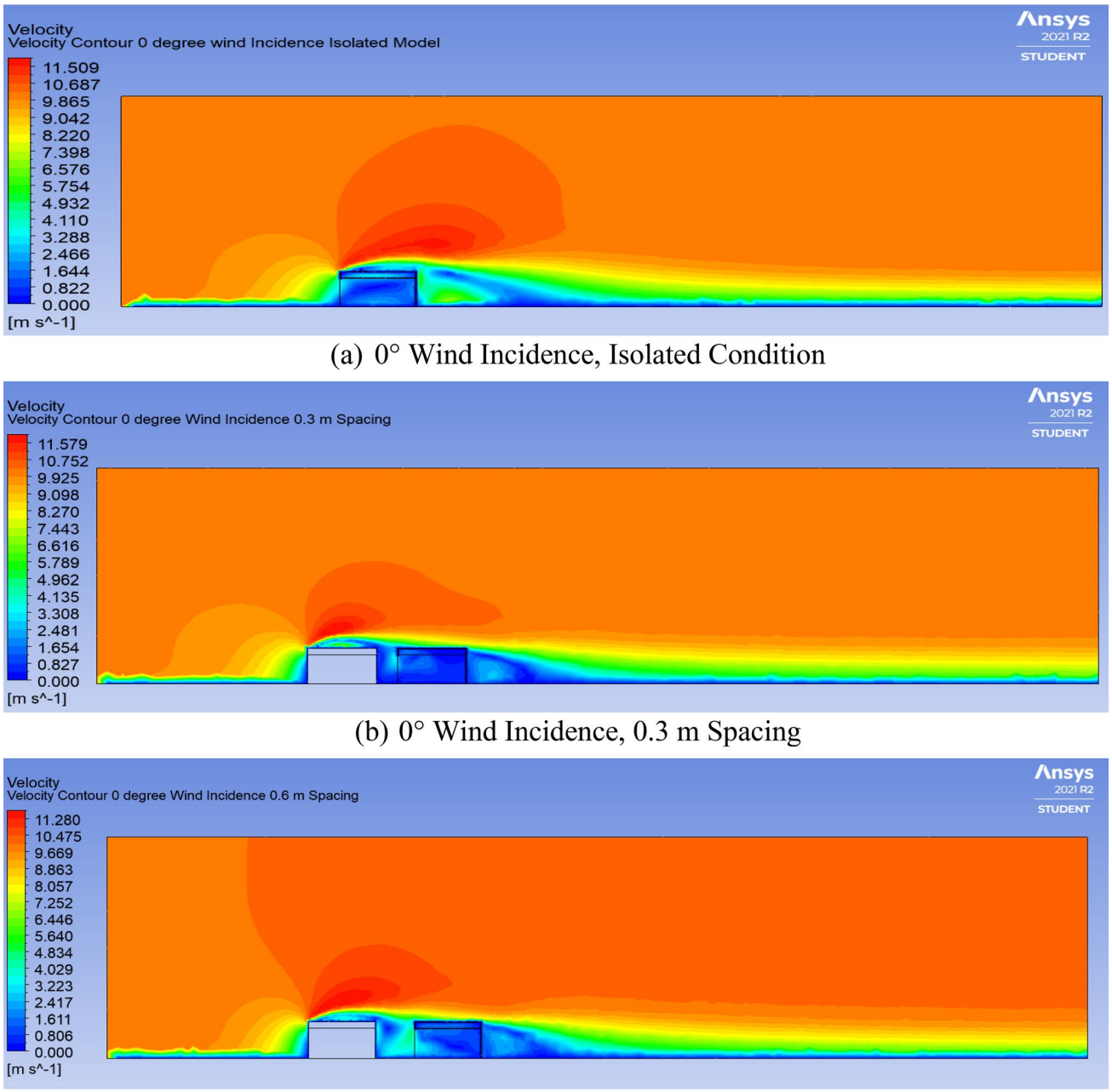
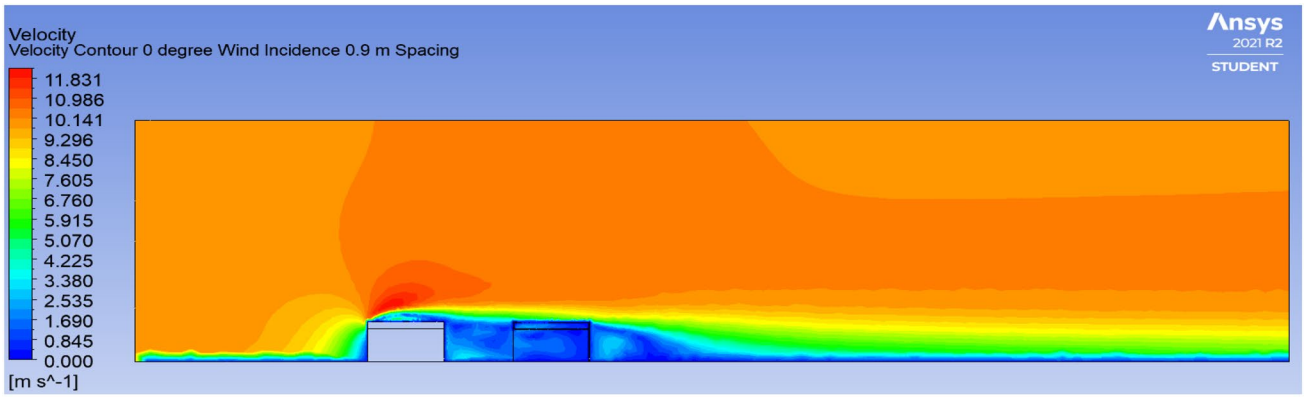
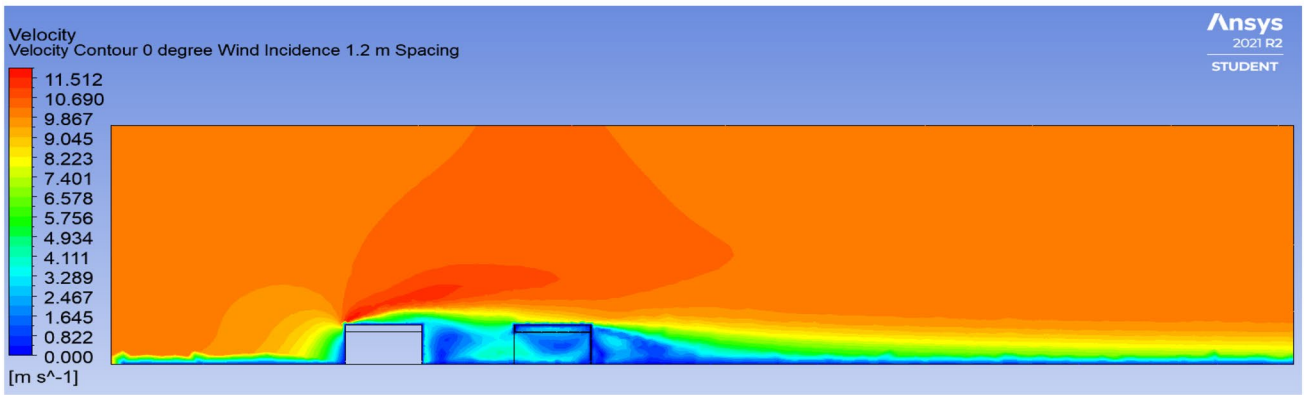


Fig. 10 Velocity contours at 0° wind incidence

red spot in the velocity contours shown in Fig. 10. When only the instrumented model is present, the maximum wind velocity is equal to 11.509 m/s. The maximum velocity is different for different spacing conditions. When the spacing between the two buildings is 0.3 m, the maximum velocity at the mid-section of the building is 11.579 m/s which is almost equal to the velocity in the isolated condition. It can also be seen that the velocity in the space between the two structures remains low and both the structures exhibit a combined bluff body action. As the spacing slightly increases to 0.6 m, the maximum velocity reduces slightly to 11.280 m/s because the wind flow rebounds after striking the instrumented building and retards the incoming flow over the interfering model. As the spacing further increases to 0.9 m, a thorough intermixing of flow in the region between the models is observed and the maximum wind velocity shows a slight increase to 11.831 m/s. The final spacing of 1.2 m corresponds to the maximum spacing in the present study, and it is observed that the maximum wind velocity is equal to 11.512 m/s which is almost equal to the maximum wind



(d) 0° Wind Incidence, 0.9 m Spacing



(e) 0° Wind Incidence, 1.2 m Spacing

Fig. 10 (continued)

velocity in the isolated condition. In this case, no increase in the wind velocity is observed since the spacing between the two structures has been increased to the point where flow is independent of the instrumented model.

From Fig. 9, IF of about 1.2 is observed for Roof B at 30° for 0.9 and 1.2 m spacings. Significant wind pressure enhancement is observed for Roof A at 60° for all spacing configurations, whereas for Roof B, a significant shielding effect is observed, especially for 0.3 and 0.6 m spacing configurations. This shows that as the spacing is increased, the shielding effect as well as the wind pressure enhancement reduces to a considerable extent. For 180° wind incidence, a rising trend is observed in the values of IF along the roof length for both Roof A and Roof B. This sharp rise is due to the reentering of wind flux into the instrumented building after striking the interfering building present on the leeward side of the structure.

From Fig. 9, it is also observed that the IF values remain fairly constant along the length of the roof except for some variations, which are seen, near points 1 and 9, which are the end regions of the roof. This is because the eddies formed over the roof escape the building surfaces near those regions. In some cases, large magnitudes of IF are also observed. This is because the wind pressure coefficient has an extremely small value at those locations for isolated cases, which results in an abnormally high interference factor. Therefore, it can be established that the study of IF alone is not sufficient to examine the interference effects of the wind. The study of IF, along with ID, is necessary to draw appropriate conclusions from the study.

Drag and Lift Coefficients

Drag and lift coefficients are important parameters of study in wind flow analysis, and they are studied along and across the wind flow direction, respectively. These coefficients represent the ratio of the forces on the structure along and across the wind flow to the product of the design wind pressure and the projected frontal/normal area of the structure. Mathematically, they are defined by Eqs. (3) and (4) [32]:

$$\text{Drag Coefficient} = C_d = \frac{F_{\text{along}}}{\frac{1}{2} \cdot \rho \cdot V^2 \cdot L_y \cdot H} \tag{3}$$

$$\text{Lift Coefficient} = C_l = \frac{F_{\text{across}}}{\frac{1}{2} \cdot \rho \cdot V^2 \cdot L_x \cdot H} \quad (4)$$

where H represents the height from the ground to the rooftop, i.e. 0.6092 m, F_{along} and F_{across} are the forces acting on the structure along and perpendicular to the wind flow direction, respectively, and L_x, L_y represent the distance between the farthest corners of the building along the x- and y-directions, respectively. The drag and lift force coefficients are considered positive along the x- and y-directions, respectively. The variation of L_x and L_y with different wind incidence angles is shown in Table 5.

The graphs in Figs. 11 and 12 represent the variation of drag and lift coefficients of the structure w.r.t different wind incidence angles. The drag coefficient for the isolated building case is nearly twice that of the 1.2 m spacing case at 0° wind incidence because of the direct striking of the wind streamline. The drag coefficient for the isolated building decreases for the intermediate angles because of the rise in effective area, whereas it increases for 90° and 180° because of similar actions of wind flow as that in 0° wind incidence. For the interference cases, the values of drag coefficients for instrumented building at 0° are lower than that of the isolated case and decrease with the decrease in spacing due to the shielding effect of the interfering building and increase at 90° for lower spacing cases due to more diversion of wind streamline by the interfering building towards the instrumented building. For the 180° case similar variation is observed as that of 0° wind incidence.

For the lift coefficients, a sharp decrease in lift coefficients by nearly 3–4 times is observed at 90°–150° wind incidence for the interference cases as compared to the isolated case. This is due to the eddies formation on the leeward side of the instrumented building leading to forces in the negative y-direction. The values of lift coefficients decrease for 0°–60° wind incidence angles for all the cases because of the increase in the effective exposed area and eddy formations which lead to forces in the negative y-direction. However, a smooth transition and similar variation of lift coefficients cannot be observed for any of the cases.

Table 5 L_x and L_y for the structure

Wind incidence angle	L_x (m)	L_y (m)
0°	1.00	0.60
30°	1.17	1.02
60°	1.02	1.17
90°	0.60	1.00
120°	1.02	1.17
150°	1.17	1.02
180°	1.00	0.60

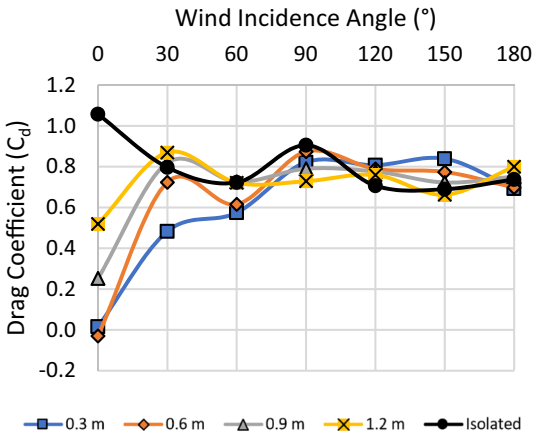


Fig. 11 Variation of drag coefficient with wind incidence angle

Conclusion

The effect of spacing configurations in interfering wind flow setup forms the base of this study. The validation of C_{pe} on the roof of a pitched roof model was carried out with the recommendations of IS 875: Part 3 (2015), and the results of CFD simulations came within satisfactory limits. Four spacing configurations, based on the width of structures, were implied to study the effects on the coefficients of pressure, drag and lift of the instrumented building and then compared with isolated wind flow conditions through interference parameters. Some of the key takeaways of this study include:

- The wind exhibits different flow behaviours for different spacing configurations which can be seen from the different maximum velocities in the velocity contours for different interference conditions in Fig. 10. The 0.3 m spacing corresponds to the combined action of the two structures, while the 1.2 m spacing configuration exhibits

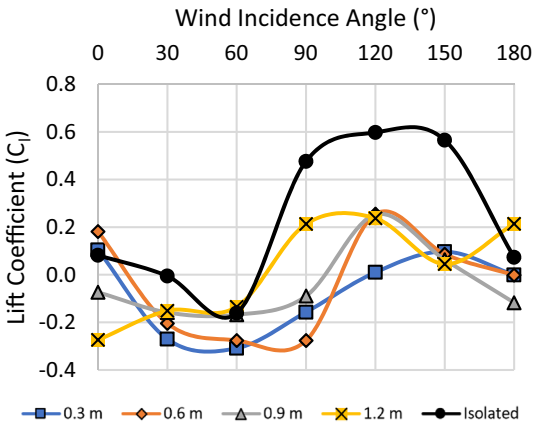


Fig. 12 Variation of lift coefficient with wind incidence angle

an independent action among the interfering and instrumented models.

- The $C_{p\ mean}$ values for both the external and internal sides of both roof faces A & B show similar variation for all the interfering conditions and the isolated case. The external sides exhibit suction pressure, whereas the internal sides experience positive pressure with larger values for the isolated case because of its direct contact with the wind streamline.
- Interference factor (IF) for the 9 points along the centre of both the roof faces has fairly similar values along the length of the roofs. The IF increases with the increase in spacings and moves towards unity, indicating a reduction in shielding and interfering effects on the instrumented building. However, the nature of variation of IF differs for each wind incidence angle.
- Extremely smaller values of C_p on the isolated model led to an abrupt rise in IF and this discrepancy is solved by calculating the interference difference (ID) of the $C_{p\ mean}$ values for each of the faces at different wind incidences.
- The values of ID were tabulated and the variation indicated reduced shielding in case of negative values and reduced wind load enhancement in case of positive values, as the spacing between the interfering building and the instrumented building was increased.
- The drag coefficients (C_d) increase with the increase in the spacing of the interfering buildings due to the reduction in shielding. The values for the interfering cases are nearly half when compared to the isolated case at 0°; however, a rise at 90° is observed in all the cases due to the direct striking of the wind streamline.
- The lift coefficients (C_l) do not follow a regular pattern of variation. However, a reduction in values for all the cases for the wind incidence period of 0°–60° and an increase in the period of 90°–150° is observed. This occurs due to the presence of eddies and the change in the effective frontal area.

It can be concluded that as the spacing between the structures was increased, a lesser shielding effect was observed. However, from the velocity contours at 0° wind incidence, it can be observed that closely spaced buildings led to stagnation of the wind between the structures, and as the spacing increases, wind velocity is increased slightly in the zone between the structures, thereby giving more natural ventilation and a negligible increase in suction on external surfaces. However, a significant rise in C_{pi} is observed, and therefore, the roof experiences increased outward push, and chances of blowing off of roof increases for higher spacing configurations. Therefore, a balanced approach needs to be adopted, by considering the factors like availability of space, the requirement of load reductions and the requirement of ventilation for the structure's function.

Among the different interference parameters based on pressure, force, moment, drag and lift coefficients, this study focuses on the interference parameters obtained through pressure coefficients. Both the IF and ID values need critical analysis for a clear observation of wind flow characteristics in interference conditions. The interference parameters obtained in this study are specific to the roof shape, its pitch angle, spacing and configuration. However, similar methodologies and careful CFD simulation techniques could be followed for other roof types and wind speed conditions. This would help in forming a comprehensive set of recommendations to be used in analyzing different interference conditions which could be adopted in widely accepted wind codes like IS 875: Part 3.

Wind-induced pressure and force coefficients and the interference parameters obtained through ANSYS CFX simulations would help decide the best possible orientations and spacings of these low-rise structures, in terms of structural soundness and serviceability under different wind flow conditions.

Funding Statement Authors would like to express their sincere gratitude to Delhi Technological University, Delhi, India, for providing funding to conduct the research work.

Declarations

Conflict of interest Authors have no conflict of interest.

References

1. P.A. Blackmore, Load reduction on flat roofs—the effect of edge profile. J. Wind Eng. Ind. Aerodyn. **29**(1), 89–98 (1988). https://doi.org/10.1016/0167-6105(88)90148-1
2. H.J. Gerhardt, C. Kramer, Effect of building geometry on roof wind loading. J. Wind Eng. Ind. Aerodyn. **43**(1), 1765–1773 (1992). https://doi.org/10.1016/0167-6105(92)90589-3
3. D. Meecham, The improved performance of hip roofs in extreme winds—A case study. J. Wind Eng. Ind. Aerodyn. **43**(1), 1717–1726 (1992). https://doi.org/10.1016/0167-6105(92)90583-V
4. K. C. Mehta, M. L. Levitan, R. E. Iverson, and J. R. McDonald, "Roof corner pressures measured in the field on a low building," *J. Wind Eng. Ind. Aerodyn.*, pp. 181–192 (1992).
5. J. A. Amin, A. K. Ahuja, "Experimental study of wind-induced pressures on buildings of various geometries," 2011. [Online]. Available: www.ijest-ng.com
6. R. Paul, S.K. Dalui, Wind effects on 'Z' plan-shaped tall building: a case study. Int. J. Adv. Struct. Eng. **8**(3), 319–335 (2016). https://doi.org/10.1007/s40091-016-0134-9
7. M. Mallick, A. Kumar, K.C. Patra, Experimental Investigation on the Wind-Induced Pressures on C-Shaped Buildings. KSCE J. Civ. Eng. **23**(8), 3535–3546 (2019). https://doi.org/10.1007/s12205-019-1929-6
8. B. Bhattacharyya, S.K. Dalui, A.K. Ahuja, Wind Induced Pressure on 'E' Plan Shaped Tall Buildings. Jordan J. Civil Eng. **8**(2), 120 (2014)

9. S. Chakraborty, K. Dalui, A.K. Ahuja, Experimental Investigation of Surface Pressure on '+' Plan Shape Tall Building. Jordan J. Civil Eng. **8**(3), 251 (2014)
10. S. Mukherjee, S. Chakraborty, S. K. Dalui, A. K. Ahuja, Wind induced pressure on 'Y' plan shape tall building. Wind Struct. Int. J. **19**(5), 523–540 (2014). https://doi.org/10.12989/was.2014.19.5.523.
11. A. K. Bairagi, A. Mukherjee, "Wind pressure and velocity pattern around 'N' plan shape tall building—A case study. Asian J. Civil Eng. **18**(8), 1241–1258, (2017) [Online]. Available: https://www.researchgate.net/publication/319620553
12. P. Sanyal, S. K. Dalui, Effects of internal angle between limbs of 'Y' plan shaped tall building under wind load. J. Build. Eng. **33** (2021). https://doi.org/10.1016/j.jobe.2020.101843.
13. P. Sanyal, S.K. Dalui, Effects of side ratio for 'Y' plan shaped tall building under wind load. Build. Simul. **14**(4), 1221–1236 (2021). https://doi.org/10.1007/s12273-020-0731-1
14. N. Gaur, R. Raj, Aerodynamic mitigation by corner modification on square model under wind loads employing CFD and wind tunnel. Ain Shams Eng. J. (2021). https://doi.org/10.1016/j.asej.2021.06.007
15. A. Kumar, R. Raj, Study of Pressure Distribution on an Irregular Octagonal Plan Oval-Shape Building Using CFD. Civil Eng. J. **7**(10) (2021). https://doi.org/10.28991/cej-2021-03091760.
16. J. Singh, A.K. Roy, Effects of roof slope and wind direction on wind pressure distribution on the roof of a square plan pyramidal low-rise building using CFD simulation. Int. J. Adv. Struct. Eng. **11**(2), 231–254 (2019). https://doi.org/10.1007/s40091-019-0227-3
17. J. Singh, A. K. Roy, CFD simulation of the wind field around pyramidal roofed single-story buildings. SN Appl. Sci. **1**(11) (2019). https://doi.org/10.1007/s42452-019-1476-2.
18. J. Singh, A.K. Roy, Wind pressure coefficients on pyramidal roof of square plan low rise double storey building. Comput. Eng. Phys. Model. **2**(1), 1–16 (2019). https://doi.org/10.22115/cepm.2019.144599.1043
19. J. Singh, A.K. Roy, Wind loads on roof of low-rise buildings. Disaster Adv. **14**(5), 83–98 (2021)
20. P. Zhou, B. Tang, P. Liu, G. Zheng, Study on Wind Load Distribution on the Surface of Dome Structure Based on CFD Numerical Simulation. J. Phys. Conf. Seri. **1176**(5). https://doi.org/10.1088/1742-6596/1176/5/052056.
21. A. Verma, A.K. Ahuja, Wind pressure distribution on rectangular plan buildings with multiple domes. Int. J. Eng. Techn. Res. (IJETR) **3**(7), 129–133 (2015)
22. A. Verma, A. K. Ahuja, Wind Pressure Distribution on Low-Rise buildings with Cylindrical Roofs. In *Implementing Innovative Ideas in Structural Engineering and Project Management*, 2015, pp. 1211–1216
23. S. Ahmad, K. Kumar, Interference effects on wind loads on low-rise hip roof buildings. Eng. Struct. **23**(12), 1577–1589 (2001). https://doi.org/10.1016/S0141-0296(01)00057-8
24. N. Gaur, R. Raj, P.K. Goyal, Interference effect on corner-configured structures with variable geometry and blockage configurations under wind loads using CFD. Asian J. Civil Eng. (2021). https://doi.org/10.1007/s42107-021-00400-0
25. S. K. Nagar, R. Raj, N. Dev, Experimental study of wind-induced pressures on tall buildings of different shapes. Wind Struct. **31**(5), 441–453 (2020). https://doi.org/10.12989/was.2020.31.5.441.
26. S. K. Nagar, R. Raj, N. Dev, Proximity effects between two plus-plan shaped high-rise buildings on mean and RMS pressure coefficients. Sci. Iran. (2021). https://doi.org/10.24200/sci.2021.55928.4484.
27. S. Pal, R. Raj, Evaluation of Wind Induced Interference Effects on Shape Remodeled Tall Buildings. Arab. J. Sci. Eng. **46**(11), 11425–11445 (2021). https://doi.org/10.1007/s13369-021-05923-x
28. T. Stathopoulos, Adverse wind loads on low buildings due to buffeting. J. Struct. Eng. **110**(10), 2374–2392 (1984). https://doi.org/10.1061/(ASCE)0733-9445(1984)110:10(2374)
29. K. Shruti, S. Rajesh, G.R. Sabareesh, P.N. Rao, Study of interference effects on Chimneys. J. Phys: Conf. Ser. **1276**(1), 012033 (2019). https://doi.org/10.1088/1742-6596/1276/1/012033
30. IS 875 (Part 3), "Design Loads (Other than Earthquake) for Buildings and Structures-Code of Practice Part 3 Wind Loads (Third Revision)." 2015.
31. "Ansys Engineering Simulation Software." 2021.
32. J.D. Holmes, *Wind Loading of Structures*, 3rd edn. (Taylor & Francis, Boca Raton, 2017)
33. *ANSYS Meshing User's Guide*, Release 13.0. Canonsburg: ANSYS, Inc., 2010.
34. A.C. Khanduri, T. Stathopoulos, C. Bédard, Wind-induced interference effects on buildings—a review of the state-of-the-art. Eng. Struct. **20**(7), 617–630 (1998). https://doi.org/10.1016/S0141-0296(97)00066-7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Check for updates

Original Article

Workplace Spirituality and Psychological Capital (PsyCap) of the Employees: A Cross-sectional Study

Global Business Review
1–19
© 2022 IMI
Reprints and permissions:
in.sagepub.com/journals-permissions-india
DOI: 10.1177/09721509221111445
journals.sagepub.com/home/gbr
SAGE

Ayatakshee Sarkar[1], Naval Garg[2] and Sasmita Palo[3]

Abstract

Grounded on the broaden-and-build theory, the study examines whether spirituality at the workplace is a significant antecedent of psychological capital. The study uses single-source data from 347 employees employed in several industries in India. Data are analyzed using partial least squares structural equation modelling (PLS-SEM) v.3.3.3. Both measurement and structural models are evaluated. The findings reveal that workplace spirituality evokes psychological capital (PsyCap) resources of hope, optimism, efficacy, and resilience. It demonstrates how employees in collectivist cultural settings embrace spirituality to conserve positive psychological resources at their workplaces. Hence, business leaders, OD practitioners need to instil and arouse spirituality-based interventions through work, climate, and organization settings. The study has many implications for business managers, HRs and OD practitioners. It highlights the significance of spiritual mind set and practices that need to be emphasized at the workplace to utilize the psychological resources embodied in their employees fully. This study is one of a kind in which the researchers have nailed down to what leads to the creation and sustainability of PsyCap at workplace.

Keywords

Workplace spirituality, PsyCap, structural equation modelling (SEM)

Introduction

One of the most significant upshots of the introduction of positive organization behaviour (POB) at the workplace is agreement over harnessing, preserving, and augmenting the employees' psychological capital (PsyCap) resources. PsyCap has emerged out of the POB literature, and it promotes the psychological well-being of the employees. It is defined as an 'individual's positive psychological state of development' (Youssef & Luthans, 2007, p. 542). Luthans et al. (2007) defined PsyCap as a constructive mental state characterized by resilience, optimism, efficacy and hope. It is seen to have a positive

[1] XLRI Jamshedpur School of Business and Human Resources, Jamshedpur, Jharkhand, India
[2] Delhi Technological University, New Delhi, Delhi, India
[3] School of Management and Labour Studies, Tata Institute of Social Sciences, Mumbai, Maharashtra, India

Corresponding author:
Ayatakshee Sarkar, XLRI Jamshedpur School of Business and Human Resources, Jamshedpur, Jharkhand, India.
E-mail: ayatakshee@gmail.com

association with several desirable workplaces human behaviour and experience, including employee well-being and satisfaction with life (Youssef-Morgan & Luthans, 2015), innovative work behaviour (Novitasari et al., 2020), creativity (Yu et al., 2019), non-violent work behaviour (Sarkar & Garg, 2020), organizational citizenship behaviours (Lifeng, 2007), employee engagement (Soni & Rastogi, 2019), workplace happiness (Kawalaya et al., 2019), work performance (Imran & Shahnawaz, 2020), and conducive organizational climate (Petersen & Youssef-Morgan, 2018). Though scores of empirical studies have reported individual and organizational outcomes of PsyCap, there are barely studies that examine its antecedents (Avey, 2014; Avey et al., 2011).

A meta-analysis on PsyCap reveals that only a few organizational antecedents like organization climate, organization justice, leader–member exchange have been explored so far (Kong et al., 2018). In simple words, there is a dearth of studies that explore what leads to the building of blissful positive psychological resources. Answering the question becomes pertinent because, at the workplace, there can be many instances when an employee's morale or productivity may dip; hence identifying factors that may positively impact PsyCap resources is necessary. The conservation of resource theory (Hobfoll, 1989) recommends protecting one's current psychological resources and acquiring new resources to remain competitive. When resources are depleted, people are more likely to feel hurt, lost and disoriented, and they may adopt harmful behaviours. Resilience, optimism, efficacy and hope are positive psychological resources if depleted, which may have negative consequences such as job burnout (Cheung et al., 2011), negative attitude (Kong et al., 2018) and detrimental health (Krasikova et al., 2015). It may also negatively influence job satisfaction, job attitude, and job performance (Kong et al., 2018).

Vilariño del Castillo and Lopez-Zafra (2022) conducted a comprehensive study on PsyCap with the help of PsycINFO and web of science (WOS) databases from 2004 to 2020. They reported that PsyCap is an essential factor in psychological health, creativity, performance, organizational climate, leadership and job satisfaction. The authors also highlighted the dire need to explore antecedents of PsyCap, especially in an organizational context. The present study tries to answer this call by examining the part of workplace spirituality in enhancing the PsyCap of employees. Founded on the broaden-and-build principle, Van Cappellen et al. (2021) suggested that spirituality provides social, emotional and cognitive resources to experience greater well-being. A few other studies also hint at the direct and indirect relationship between workplace spirituality and PsyCap. Workplace spirituality significantly influences an individual's mental, moral and physical makeup (Pawar, 2017; Petchsawang & McLean, 2017). Fox et al. (2018) noted that cognitive factors like spirituality might serve as a potential influencer on PsyCap. Similarly, authentic leadership promotes PsyCap and creativity (Rego et al., 2012). Prior studies have reported many positive consequences of workplace spirituality that build up positive psychological resources. Workplace spirituality increases inner strengths (Sepahvand, 2019), brings a positive transformation in the outlook (Mahipalan & Sheena, 2019), nurtures relationships (Suárez, 2015) and promotes non-violent work behaviour (Sarkar & Garg, 2020).

Although researchers have suggested various potential antecedents' variables of PsyCap, the present study is particularly interested in exploring workplace spirituality for two reasons. First, workplace spirituality is a holistic alignment of employees' bodies, minds and spirits (Pandey, 2017). Thus, it is more likely to have a more significant and profound impact than any other organizational factor. Second, workplace spirituality honours the inner self of an employee. It focuses on the inner self and the outer environment of workplace (Lips-Wiersma & Morris, 2009). Moreover, an employee spiritually anchored to the subtle inner self is more poised to conserve positive psychological resources that help build positive emotions at the workplace. Also, the present study answers the call of the prior researchers that attempted to develop theoretical associations between spirituality and PsyCap (Fox et al., 2018; Golparvar & Jafari, 2013; Jena & Pradhan, 2015; Paul & Saha, 2016). The uniqueness of the

present study lies in two novel contributions. First, the study explores Indian dimensions of workplace spirituality (meaningful life, meditative work, collaborative problem solving, authenticity and *lokasangraha*). Previous studies have mainly relied on Western conceptualization of spirituality at work. Second, this study is based on higher-order reflective–formative conceptualization of workplace spirituality. Such a method allows deeper insights into the proposed relationships. Based on the above theoretical framework, the present study aims to investigate the following research question:

- Does workplace spirituality positively impact PsyCap?

Literature Review

Psychological Capital and its Dimensions

The foundation of PsyCap is found in Seligman's (2004) 'Authentic Happiness.' Here it states that 'experience of flow at work' builds up psychological resources. Stajkovic (2003) also revered the importance of PsyCap for work motivation. Famous positive psychologist Csikszentmihalyi (2003, p. 26) stated that

> it is developed through a pattern of investment of psychic resources that results in obtaining experiential rewards from the present moment while also increasing the likelihood of future benefit. It's about the state of the components of your inner life. When you add up the components, experiences, and capital, it makes up the value.

After these initial references, Luthans et al. (2007) further enhanced the work of PsyCap. Their definition and constructs established PsyCap in the literature of POB and OB.

Following Luthans et al. (2004), PsyCap is beyond 'human and social capital'. While 'economic capital' focused on 'what you have?', 'human capital' on 'what you know?' and 'social capital' on 'whom you know?', 'Psychological Capital' focuses on 'Who you are?' and 'what you can become?' (Luthans & Avolio, 2003; Luthans et al., 2006). That is, its primary emphasis is on psychological strengths. Thus, PsyCap can be defined as

> an individual's positive psychological state of development characterized by: (1) having confidence (self-efficacy) to take on and put in the necessary effort to succeed at challenging tasks; (2) making a positive attribution (optimism) about succeeding now and in the future; (3) persevering toward goals, and when necessary, redirecting paths to goals (hope) in order to succeed; and (4) when beset by problems and adversity, sustaining and bouncing back and even beyond (resilience) to attain success. (Luthans et al., 2007, p. 3)

The integrated four psychological constructs form 'a higher-order latent construct' known as 'PsyCap' (Luthans et al., 2007).

As discussed, PsyCap constituents are 'self-efficacy, hope, optimism and resilience' (Luthans et al., 2007). Snyder et al. (1991) define the first element of hope as 'a positive motivational state supported by cognitive processes of (1) agency (goal-directed motivational energy) and (2) pathways (planning multiple paths to meet goals)' (p. 287). 'Hope' is a crucial psychological asset as it inspires one to carry on the journey regardless of the current situation. The 'hopeful person' tries out innovative means and ways to cope with problematic situations, whereas a 'low hope' person finds it extremely challenging to cope due to 'fear of failure' (Snyder, 2000, 2002; Snyder et al., 1996). The second dimension is

self-efficacy, defined as 'people's beliefs in their capabilities to produce desired effects by their actions' (Bandura, 1997, p. vii). In other words, Maddux (2009) described 'self-efficacy' as 'what I believe I can do with my skills under certain conditions' (p. 336). The third construct, 'optimism', is defined as 'positive generalized outcome expectancies' (Scheier Carver, 1985, p. 219). A person with an optimistic mind set attributes failures to external cause and circumstances, whereas pessimists attribute to personal reason and blame themselves for probable cause (Seligman, 1998).The last element of PsyCap is 'resilience', which according to Luthans (2002), is 'the developable capacity to rebound or bounce back from adversity, conflict, and failure or even positive events, progress and increased responsibility' (p. 702). 'A resilient person' has an 'exceptional capacity to bounce back' from the adverse and stressful situations positively. Such a person also exhibits remarkable attributes to 'bounce back' from failure and loss (Masten & Reed, 2002).

The psychometric properties of these four constructs of PsyCap are well researched and documented by previous researchers. Several authors, including Bandura (1997), Snyder (2002) and Luthans et al. (2007), theoretically and conceptually elaborated that resilience, optimism, efficacy and hope are distinct constructs. Other scholars (e.g., Bryant & Cvengros, 2004; Carifio & Rhodes, 2002; Luthans et al., 2007; Youssef & Luthans, 2007) provided empirical support for discriminant validity among four dimensions of PsyCap.

Workplace Spirituality

Researchers have no agreement for a unanimous 'definition' of 'workplace spirituality'. Ashmos and Duchon (2000) defined 'workplace spirituality' as 'the recognition of an inner life nourished by meaningful work which takes place in the context of community.' Workplace spirituality is 'a framework of organizational values evidenced in the culture that promotes employees' experience of transcendence through the work process, facilitating their sense of being connected to others in a way that provides feelings of completeness and joy' (Giacalone & Jurkiewicz, 2003). Zohar and Marshall (2000) defined 'workplace spirituality' in the form of 'spiritual intelligence'. It is the intelligence through which one understands their 'highest meanings, purposes, and motivations.' In a meta-analysis study, Karakas (2010) reported '70 definitions of workplace spirituality' in the available literature of workplace spirituality. In simple words, it refers to the experience and realization of spirit at the workplace (Garg et al., 2019).

Silver lining amid confusion for uniform definition is comprehensive agreement over multidimensionality of workplace spirituality (Dehler & Welsh, 2003). Although it is considered a multidimensional construct, scholars recommend different dimensions of workplace spirituality owing to its cultural sensitivity. Mitroff and Denton (1999) suggested three dimensions: meaning in the work–life, virtuous interpersonal relationship and harmony with organizational values. Ashmos and Duchon (2000) also advocated its three elements: inner life, meaningfulness at work and sense of belonging. Sheep (2006) recommended four dimensions: inner life, transcendence, self-workplace integration, and meaningful and purposeful work. Since the present study is conducted in Indian settings, the Pandey et al. (2019) scale is preferred. The said scale has five dimensions of Indian workplace spirituality: meaningful work, meditative work, collaborative problem solving, authenticity, and *lokasangraha*.

Meaningfulness at the workplace ensures perfect synchronization of one's 'life purpose with its work profile and work accomplishment' (Pandey et al., 2019). A work assignment following one's talent, interest, abilities and values also allow an employee to cherish meaningfulness at the workplace. People often leave their present job to pursue more meaningful work (Burack, 1999). Meaningfulness is one of

the most significant components of workplace spirituality (Dirkx, 2001, 2013). It acknowledges employees' inner self (Pardasani et al., 2014). The employer has a prime responsibility to make employees' stay meaningful (Adawiyah & Pramuka, 2017). Meditative work is a state of being one with the work (McCormick, 1994). It calls for a 100% devotion and dedication to the work. It aligns one's mind, body and soul at the workplace. Authenticity is another dimension of workplace spirituality that recommends truthfulness, honesty, non-stealing, non-violence, and accountability with self and others (Garg & Gera, 2019; Mehta et al., 2022). It also encourages the oneness of thoughts, words and actions. Collaborative problem solving develops a feeling of belongingness and commitment among the group members. It nurtures the belief that each group member can rely on others during challenging and tiring lifetimes. Their interdependence and interconnectedness create a 'sense of community' among them. The endmost dimension of workplace spirituality is *lokasangraha* that means 'coming together for the welfare of the others' (Pandey et al., 2019). Contemporary employers are trying to arrange avenues to fulfil their employees' physical, emotional, and spiritual needs at the workplace (Bhunia & Das, 2012). Moreover, *lokasangraha*, that is, collaborative efforts for the well-being of society, is one of the revered ways of flourishing at the workplace (Garg et al., 2021; Pandey et al., 2019; Pradhan & Jena, 2017; Saxena et al., 2020; Sharma et al., 2022).

Workplace Spirituality and Psychological Capital

Plenty of studies delineate the inter-linkages between spirituality and 'positive psychological constructs' (Singh et al., 2016). Spirituality also leads to gratitude, faith, benevolence and harmony among all human beings. Mahoney et al. (1999) found that 'spiritual individuals' were more hopeful and had greater purpose and meaning in life. A person who believes in spirituality and spiritual practices are seen to be less stressed and with no depression symptoms. Moreover, it also affects the components of three *gunas* in shaping one's personality (Singh et al., 2016).

The proposed relation between workplace spirituality and PsyCap is theoretically premised in Fredrickson's broaden-and-build theory of positive emotions. According to Fredrickson, positive emotions 'broaden people's momentary thought–action repertoires and build their enduring resources' (Fredrickson, 2001, p. 147). It is well-researched that spirituality leads to positive thoughts, emotions, and behaviour (Aboobaker et al, 2019; Garg & Mehak, 2021; Vasconcelos, 2018). And it is being proposed that spirituality-induced positive emotions will build psychological resources for the employees. This study suggests that these psychological resources are PsyCap, a collective term used for hope, optimism, self-efficacy and resilience. Limited studies have investigated the direct linkage between workplace spirituality and PsyCap. Pedreira and Mónico (2013) reported a positive association between these variables in an organizational context. Another study revealed similar positive correlation results between workplace spirituality and PsyCap of the manufacturing sector employees (Jena & Pradhan, 2015). Several other studies provide empirical evidence of the connection between spirituality and PsyCap in the workplace, such as nursing, coping, gerontology and substance abuse. Several organizational factors can substantiate workplace spirituality leading to hopefulness among employees (Curry et al., 1997), increased self-esteem (Barnum et al., 1998; Pirkola et al., 2016) and enhanced work satisfaction (Schwartz, 2000). In an organizational climate that promotes hope, employees are better able to deal with workplace stressors (Edwards & Cooper, 1988; Simmons & Nelson, 2001).

Apart from workplace spirituality, individual spiritual orientation is also reported to link with PsyCap directly. Snyder (2002) reported a close association between spirituality and hope. In studying the rural

Afro-American females who have breast cancer, Powe (1997) acknowledged a relationship between a sense of fatalism and spirituality. She states that spiritual people had higher hopes and better copying abilities than people with a lesser spiritual orientation. Ironson (2007) states, 'Spirituality gives people a sense of peace and optimism. It gives them hope' (p. 430). Many social work authors describe how spirituality inspires and permeates their work for social justice. Spirituality serves as a source of inspiration and strength for social activism that promotes social justice (Nash & Stewart, 2005).

Among senior adults, spirituality promoted resilience (Manning, 2013). Religiousness and spirituality are found to strongly influence the perception of individuals under stress, like trauma victims and survivors of Nazi concentration camps. Behaviours of such people correspond with the resilient attributes of these survivors. (Peres et al., 2007). According to Kim and Esquivel (2011), spirituality is a source of resilience for adolescents. Workplace spirituality also promotes coping with adversaries. Above discussed studies investigated the association between spirituality and a few facets of PsyCap. It would be interesting to examine the role of workplace spirituality on PsyCap in an organizational context, knowing the fact that PsyCap contributes significantly to employee's commitment to Indian organizations (Sahoo & Sia, 2015). Based on the above arguments, the following hypothesis is formulated.

H1: Workplace spirituality has a significant positive impact on PsyCap.

Workplace Spirituality as a Second-order Construct

The current article proposes workplace spirituality as a higher (second-order) reflective–formative construct. Moreover, its five dimensions (meaningfulness, meditative work, collaborative problem-solving, authenticity and *lokasangraha*) are taken as first (lower) order reflective constructs. These five factors have their independent conceptual meanings reflected in the measures; therefore, they are referred to as 'reflective first-order constructs'. Researchers have proposed a higher-order construct to reduce the number of hypothesized relationships and make the model easy to understand and interpret (Sharma et al., 2021; Thien, 2020). In addition, it also lessens the multi-collinearity issues (Sarstedt et al., 2019). To sum it up, the higher-order constructs (workplace spirituality) are measured by their first-order constructs. Moreover, the first-order reflective or formative constructs are measured by their measurement items.

Methodology

Sample

The study is based on data from a cross-section survey. In order to gain entry to the organisation, the HR heads were first contacted through emails. Initially nine organizations were approached across sectors seeking permission for the study; however permissions were approved only by three of them. Once permission to enter the organisation was granted, brief meeting was set up individually with all the department heads for about 15–25 minutes, as per their convenience. It was only after the survey was understood and approved by the department head, their team was called upon to engage in the second round of meeting, briefing them on the survey, meaning of the constructs and some key instructions to fill up the survey. In the meeting, each respondent were assured of their confidentiality with regard to

their identity, hence they were emphasized to be as honest as possible with their responses. Out of total 430 surveys distributed, only 412 were returned, out of which 347 were considered. The respondents in the study are professionals across functions such as production, research and development, marketing, human resource management, finance and accounting, legal, and information technology. Some of their designations are consultants, supervisors, managers, senior managers, legal officers, claims analysts, vigilance officers across several industries such as banking and financial services, oil and gas, education, IT, and services working in Mumbai and Bangalore region. The composition of sample respondents is as follows: 70.1 % male and 29.9 % female, and their age groups are up to 34 years (55.8%), 35–44 years (27.7%), above 44 years (16.4%). The mean age of the respondents is 35.06 years. The average experience in the current organization is 7.04 years, and the mean total experience is 12.03 years. Maximum variation sampling, that is, a purposive sampling technique, is used to capture a range of perspectives related to the constructs. The central notion of the sampling is to look at the respondents from all available angles, that is, a wide range of backgrounds, behaviours (Etikan et al., 2016).

Measure

PsyCap is measured using a scale developed and validated by Luthans et al. (2007). The scale comprises 12 statements that measure four dimensions of PsyCap, that is, resilience, optimism, efficacy, hope. It is a 6-point rating scale ranging from 1 (strongly disagree) to 6 (strongly agree). A few items of the scale are 'I usually take stressful things at work in stride,' 'I can think of many ways to reach my current work goals,' and 'I feel confident to present information to a group of people.'

Workplace spirituality is assessed with the help of a scale proposed by Pandey et al. (2019). It is an 11 items scale that captures five dimensions of workplace spirituality: meaningfulness, meditative work, collaborative problem solving, authenticity, and *lokasangraha*. It is a 5-point rating scale ranging from 1 (strongly disagree) to 5 (strongly agree). A few items of the scale are 'my job helps me to understand my life's purpose,' 'working here makes my life meaningful,' 'peoples' actions here are aligned with their words,' and 'people here own up to mistakes with others in the group.' Since researchers hail workplace spirituality as a cultural-sensitive term, thus the present scale is apt for the study, as it is developed and validated among Indian respondents.

Data Analysis

PLS-SEM is used to explore the proposed relationship between workplace spirituality and PsyCap. PLS-SEM has been preferred over other statistical analysis tools as it has a unique ability to deal with 'higher-order reflective–formative constructs' (Cheah et al., 2020). It makes the PLS-SEM model more comprehensive for data analysis. In the present study, type two higher-order (reflective–formative) constructs are proposed (Sarstedt et al., 2019). Hair et al. (2019) suggested that the structural model has been assessed after exploring the measurement model.

Results

Normality of Data and Common Method Variance (CMV)

The Kolmogorov–Smirnov test and the Shapiro–Wilk test are used to check normality of the data. The significance values in both tests are above 0.05, which means tests are not significant. And hence, it could be concluded that data is normal. Data is collected from a single source using self-reported measures; thus, CMV is possible (Mackenzie & Podsakoff, 2012). The current study has minimized the risk of CMV by applying both procedural design and statistical tactics (Reio, 2010). First, pre-validated scales with different rating scales measure workplace spirituality and PsyCap. Second, respondents are promised confidentiality and anonymity of the data. They are also assured that the collected data will be used for academic purposes and not be shared with anyone. Such measures encourage respondents to respond freely and reliably (Podsakoff et al., 2012). Kock (2015) suggested two statistical control methods, that is, Harman's 1-factor test and collinearity tests, to assess the significance of CMV. In Harman's test, more than one factor emerges, and the variance explained by the largest factor is 31%. It negates the significance of CMV for the present study. Also, the collinearity tests results indicate values between 1.0 and 2.7 which is lesser than the threshold value of 3.3. Thus, these results suggest that CMV is not a significant issue in the study.

Assessment of Reflective Constructs

The measurement model evaluation aims at investigating the reliability and validity of manifest variables using the values of outer loading, average variance extracted (AVE), composite reliability (CR), and discriminant validity (Hair et al., 2019). Table 1 reports that the outer loadings of the first-order reflective construct (workplace spirituality and PsyCap items) are above the minimum threshold value of 0.70 (Henseler et al., 2009). It confirms reliability or consistency. Also, the CR values of all variables are greater than 0.7, which concludes a high internal consistency. All AVE values are greater than 0.5, which assures convergent validity. Similarly, Rho_A values are greater than the threshold value of 0.70, confirming composite reliability (refer to Table 1). Heterotrait–Monotrait ratio (HTMT) criterion is used to evaluate discriminant validity. HTMT is a more recent, established, and effective measure than Fornell–Larcker's criterion (Shiva et al., 2020). HTMT values should be less than 0.85 and 0.9 (permissible value) (Gold et al., 2001; Henseler et al., 2015). Table 2 reports that HTMT values of most of the variables except hope and resilience are acceptable. Since the confidence intervals for these constructs are well within limits, a liberal HTMT approach may be applied (Shiva et al., 2020).

Assessment of Formative Constructs

Workplace spirituality is proposed as a higher-order reflective formative construct. The variance inflation factor (VIF) and the significance of the indicator are used to evaluate the formative measure. Table 3 shows that all VIF values are less than three, which serves as a threshold value (Peter et al., 2007). It suggests that multi-collinearity is not a significant issue with five dimensions of workplace spirituality. A bootstrap on 5000 subsamples is also run to assess the significance of the weights. Table 3 indicates that the results are significant at the $p < 0.01$ level. It confirms the significance of the formative construct to the creation of a higher-order reflective–formative construct.

Table 1. Measurement Model Evaluation.

Construct	Item	λ	AVE	CR	Rho_A
ML	ML1	0.882	0.810	0.928	0.884
	ML2	0.908			
	ML3	0.910			
MW	MW4	0.917	0.845	0.916	0.817
	MW5	0.922			
CW	CW6	1			
AU	AU7	0.92	0.852	0.92	0.828
	AU8	0.927			
LO	LO9	0.900	0.622	0.823	0.757
	LO10	0.901			
	LO11	0.795			
O	O1	0.922	0.801	0.923	0.878
	O2	0.894			
	O3	0.867			
HO	HO4	0.848	0.76	0.927	0.895
	HO5	0.873			
	HO6	0.904			
	HO7	0.862			
RE	RE8	0.849	0.744	0.897	0.829
	RE9	0.878			
	RE10	0.861			
SE	SE11	0.940	0.875	0.933	0.860
	SE12	0.931			

Source: The authors.

Note: WS: Workplace spirituality (ML: meaningfulness. MW: meditative work, CW: collaborative problem solving; AU: authenticity, LO: *lokasangraha*); PC: psychological capital (O: optimism, HO: hope: RE: resilience; SE: self-efficacy), λ: outer loading, AVE: average variance explained, CR: composite reliability.

Table 2. Discriminant Validity (Heterotrait–Monotrait ratio criterion).

Variable	AU	CW	O	HO	LO	ML	MW	RE	SE
AU									
CW	0.129								
O	0.544	0.221							
HO	0.600	0.211	0.902						
LO	0.736	0.229	0.612	0.536					
ML	0.689	0.264	0.667	0.689	0.770				
MW	0.687	0.252	0.792	0.785	0.707	0.845			
RE	0.473	0.220	0.850	0.919	0.439	0.548	0.725		
SE	0.617	0.206	0.809	0.853	0.484	0.618	0.734	0.851	

Source: The authors.

Note: WS: Workplace spirituality (ML: meaningfulness. MW: meditative Work, CW: collaborative problem solving; AU: authenticity, LO: *Lokasangraha*); PC: psychological capital (O; optimism, HO: hope: RE: resilience; SE: self-efficacy).

Table 3. Assessment of Formative Constructs.

Variable	FI	VIF	*t*-Value	*p*-Value
Workplace spirituality	ML1	2.225		
	ML2	2.270	46.657	0.000*
	ML3	2.729		
	MW4	1.906		
	MW5	1.931	45.549	0.000*
	CW6	1		
	AU7	1.987		
	AU8	1.912	45.888	0.000*
	LO9	1.832		
	L010	1.739	44.920	0.000*
	LO11	1.840		

Source: Primary data.

Note: *$p < 0.01$, FI: formative indicators; VIF: variance inflation factor; ML: meaningfulness. MW: meditative work, CW: collaborative problem solving, AU: authenticity; LO: *Lokasangraha*.

Assessment of the Structural Model

The relationship between second-order constructs, that is, workplace spirituality and PsyCap, is examined using the structural model (Hair et al., 2019). Figure 1 indicates workplace spirituality and PsyCap are second-order constructs, with workplace spirituality having formative assessments. It is found that both the outer and the inner weights are significant at 1% and different from zero (Henseler, 2017). It is also reported that the path coefficients between workplace spirituality and PsyCap lie between −1 and +1. The goodness of fit criterion is assessed using standardized root mean square (SRMR). SRMR is the most vital criterion to check the model's goodness (Hair et al., 2020). The current model has an SRMR value of 0.046, which is far below the threshold of 0.08 (Henseler et al., 2016). It confirms the goodness of the model. Table 4 reflects that workplace spirituality is a significant predictor of PsyCap with $\beta = 0.737 (p < 0.001)$, thereby supporting the hypothesis of the study. The coefficient of determination (R^2) is 46.1%, which means workplace spirituality could explain 51.1% variations in PsyCap. Since the relationship between second-order constructs (overall workplace spirituality and overall PsyCap) is investigated using the PLS model, multiple regression is used to evaluate the linkage between different dimensions of workplace spirituality and PsyCap. Regression results indicate a significant regression model ($F = 66.245$, $p < 0.05$). All five dimensions of workplace spirituality reported significant standardized regression coefficients. The values of regression coefficients are 0.324, 0.555, 0.414, 0.472, and 0.376 for meaningful work, meditative work, collaborative problem solving, authenticity, and *lokasangraha*.

Discussion

Theoretically postulated on Fredrickson's (2001) 'broaden-and-build theory', the present study proposes that workplace spirituality positively predicts employee PsyCap resources. The study contributes to the literature by examining workplace spirituality as a resource building and investing

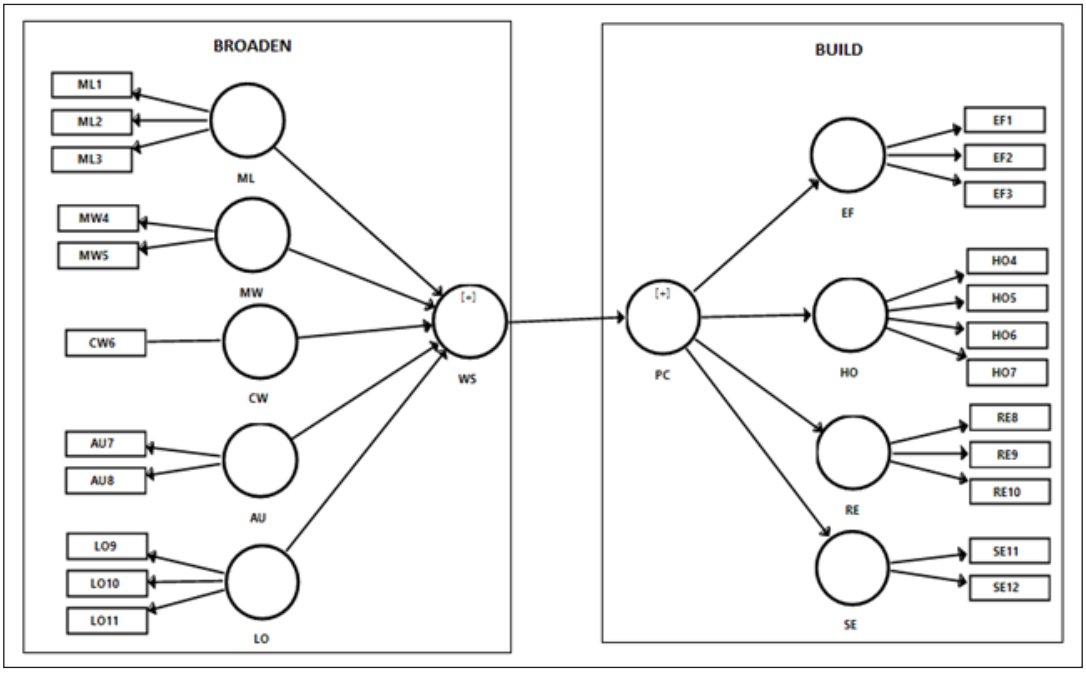


Figure 1. PLS-SEM Model.

Source: Primary data.

Note: ML: Meaningful life, MW: Meditative work, CW: Collaborative problem solving, AU: Authenticity, LO: *Lokasangraha*, OP: Optimism, HO: Hope, RE: Resilience, SE: Self-efficacy.

Table 4. Assessment of the Structural Model.

Hypothesis	Path Relationships	Std. Beta	Sample Mean (M)	*t*-Values	CI 2.5%	CI 97.5%	Decisions
H1	WS→PC	0.737	0.730	14.292***	0.569	0.755	Supported

Source: Primary data.

Note: Path co-efficient (*p < 0.01, **p < 0.05, ***p < 0.001); workplace spirituality; PC: psychological capital.

concept towards sustaining PsyCap resources. Exploring the association between workplace spirituality and PsyCap is still nascent (Jena & Pradhan, 2015). Several authors (e.g., Jena & Pradhan, 2015; Luthans et al., 2007; Paul & Saha, 2016) have attempted to understand spirituality as a potential PsyCap tool. Therefore, it is pertinent to examine the linkage between workplace spirituality and PsyCap. PsyCap as a positive psychological construct is good for employees' well-being and performance. Nevertheless, the current study has emphasized what triggers the formation and sustainability of these positive psychological states.

The study reports that workplace spirituality is a strong predictor of PsyCap of the employees. Besides overall workplace spirituality, each spirituality dimensions show potentiality to determine the

PsyCap of the employees. Previous researchers have already recommended spirituality dimensions like meaningful work. Meaningful work is also a key factor in motivating employees. Several other authors like Lepisto and Pratt (2017), Michaelson et al. (2014), Rosso et al. (2010) have also focused on meaningful work through individual and organizational lenses. On the other hand, meditative experiences at work is being completely immersed into the work for the greater good of the society (McCormick, 1994; Pandey et al., 2019). Both meaningful and meditative work has been drawn from the Indian tradition of one's *swadharma* (Pandey, 2017). Study reports that when employees are engaged in meaningful and meditative work, they are more likely to remain resilient, hopeful and optimistic (Lepisto & Pratt, 2017; Michaelson et al., 2014).

Similarly, the study advocates that *lokasangraha* (coming together for the welfare of society) also promotes psychological resources of the employees. Other's orientation provides a broader vision and meaning to one's life (Saxena et al., 2020). A broad vision of universal welfare gives one hope and optimism of a brighter future. *Lokasangraha* creates a feeling of connections at the workplace, thereby imbibing harmony, optimism and hopes in relationships. Indulgence in work related to others' betterment develops self-confidence and self-efficacy. Besides, literature from other disciplines such as counselling psychology, gerontology and nursing brings out 'spirituality' as a means of healing, resilience, hope, confidence, maintaining a balanced state of mind and optimism. Thereby, they also support the findings of the study.

In addition, authenticity and collaborative problem solving also contributes significantly to the manifested dimension. Authenticity is the core aspect of the human self that every organisation looks out for especially during the selection process. It is about coherence of people's behaviour with the shared beliefs and values they possess (Pareek, 2002). It is demonstrated in the form of several decision making and problem solving activities. On the other hand, collaborative problem solving is a cooperation technique that fosters optimism, hope and self-efficacy not only with the peers but also with the opponents (Weber, 2001).

Theoretical Implications

The current study makes at least five valuable contributions to human resource management and positive psychology literature. First, this study extends 'the conservation of resource (COR) theory' by suggesting a means to conserve and augment psychological resources. It proposes to encourage COR theorists to examine the role of spirituality in psychological resources. Second, the study's findings support an organization's resource-based view, which reinstates that besides investing in financial resources, organizations should also invest in social and psychological resources. It may promote studies investigating how workplace spirituality leads to competitive advantage by ensuring employees' valuable, rare, inimitable and non-substitutable PsyCap. Third, Rousseau (2014) propagates a system theory of spirituality in which the author introduced a theoretical concept of 'spiritual welfare.' 'Spirituality can be linked to psychological factors that increase stress tolerance and immune responses, and to dietary and other behaviours that reduce medical risks' (Rousseau, 2014, p. 483). The present study offers empirical support to the theoretical concept of spiritual welfare. Here, various dimensions of workplace spirituality like meaningful work, *lokasangraha*, etc., are reported to perform spiritual welfare functions. Fourth, workplace spirituality is often criticized for the lack of empirical studies. The limited attention to spirituality at the workplace has restricted the development and application of

theories of workplace spirituality (Pawar, 2009; Sarkar & Garg, 2021). Spirituality's 'inclusion in established research streams is long overdue' (Sheep, 2006, p. 371). This study reports an organizational utility of spirituality and therefore it contributes significantly to the existing knowledge of workplace spirituality.

Moreover, the present study provides empirical evidence to Fredrickson's (2001) 'broaden-and-build theory of positive emotions' in non-Western settings. The broaden-and-build view was developed and validated in Western settings. The study's findings suggest that workplace spirituality builds psychological resources in Indian respondents. It helps in cross-cultural validation of the theory.

Managerial Implications

The study's findings offer significant workable suggestions for the managers and employers. Since PsyCap directly influences the performance of the employees (Imran & Shahnawaz, 2020), the study results establish workplace spirituality as a significant predictor of PsyCap. Thus, HR managers may try to develop a spiritual organizational culture to build and restore the 'psychological resources' of the employees. The results suggest that meaningful and meditative work assignment leads to the development of PsyCap. Thus, managers may opt for job redesign or reengineering in light of the job characteristics model to make job profiles of the employees more meaningful and lively. A positive work environment that includes sharing the company's vision, employee participation in decision-making, open communication, the delegation of the authority, etc., could help employees find meaning in their work (Daniel, 2015). Further, collaborative problem solving could be ensured through group assignments, group projects, team-based incentives, rewards, etc. The results also report a positive linkage between *lokasangraha* and PsyCap. Thus, the manager should convince the employees that they perform organizational work and serve society through their goods and services. Similarly, contemporary companies are adopting novel initiatives such as cause-related marketing, green HRM and activities under corporate social responsibility, to tap the benefits of 'desire to help others.' Further, organizations may resort to specific spiritual practices such as meditation, yoga, *pranayama, vipassana,* self-introspection and mindfulness, to develop the spiritual orientation of the employees. Few progressive companies are going the extra mile to cherish the benefits of spirituality at the workplace. For example, Google and YouTube have started meditation rooms in their office premises. Microsoft has begun a program of online prayer for its employees. Also, spiritual leadership theory highlights the importance of ethical and spiritual leadership to disseminate spiritual values among employees (Fry, 2003). Thus, findings of this study, previous studies and observation of current trends in business organizations strongly pitch for the adoption of spiritual values and practices at the workplace.

Limitations, Scope for Future Studies, and Conclusion

Although the current study reports significant findings, it has certain limitations too. First, CMV could be an issue as the variables are collected from single source. Although appropriate remedies have been adopted to deal with CMV, repeated cross-sectional, longitudinal, or multisource data could help generalize the results. Second, the present study is based on quantitative data only. Future researchers may use qualitative techniques like theme analysis, framework analysis to highlight the underlying

dimensions of workplace spirituality in Indian organizations. These underlying factors could further be subjected to quantitative analysis to investigate their association with PsyCap. Third, the study is based on a sample of 347 only. Larger sample size may help with more robust and acceptable results. Also, cross-cultural researches are essential to generalize the results for other cultures. Future researchers may also examine other organizational factors like leadership, job satisfaction, etc., for any possible linkage with PsyCap. In Conclusion, the study's findings have a broader scope and are not limited to a workplace context. Maintaining a positive psychological state is relevant to all spheres and across all domains of life where one looks forward to creating a better life and accumulating good fortune.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Ayatakshee Sarkar https://orcid.org/0000-0002-9620-4395

Naval Garg https://orcid.org/0000-0001-6694-9663

References

Aboobaker, N., Edward, M., & Zakkariya, K. A. (2019). Workplace spirituality, employee wellbeing and intention to stay: A multi-group analysis of teachers' career choice. *International Journal of Educational Management*, *33*(1), 28–44.

Adawiyah, W. R., & Pramuka, B. A. (2017). Scaling the notion of Islamic spirituality in the workplace. *Journal of Management Development*, *36*(7), 877–898.

Ashmos, D. P., & Duchon, D. (2000). Spirituality at work: A conceptualization and measure. *Journal of Management Inquiry*, *9*(2), 134–145.

Avey, J. B. (2014). The left side of psychological capital: New evidence on the antecedents of PsyCap. *Journal of Leadership & Organizational Studies*, *21*(2), 141–149.

Avey, J. B., Reichard, R. J., Luthans, F., & Mhatre, K. H. (2011). Meta-analysis of the impact of positive psychological capital on employee attitudes, behaviors, and performance. *Human Resource Development Quarterly*, *22*(2), 127–152.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. H. Freeman.

Bhunia, A., & Das, S. A. (2012). Explore the impact of workplace spirituality on motivations for earnings management-an empirical analysis. *International Journal of Scientific and Research Publications*, *2*(2), 153–161.

Blumenthal, J. A., Babyak, M. A., Ironson, G., et al. (2007). Spirituality, religion, and clinical outcomes in patients recovering from an acute myocardial infarction. *Psychosomatic Medicine*, *69*(6), 501–508.

Bryant, F. B., & Cvengros, J. A. (2004). Distinguishing hope and optimism: Two sides of a coin, or two separate coins? *Journal of Social and Clinical Psychology*, *23*(2), 273–302.

Burack, E. H. (1999). Spirituality in the workplace. *Journal of Organizational Change Management*, *12*(4), 280–292.

Carifio, J., & Rhodes, L. (2002). Construct validities and the empirical relationships between optimism, hope, self-efficacy, and locus of control. *Work*, *19*(2), 125–136.

Cheah, J. H., Roldán, J. L., Ciavolino, E., Ting, H., & Ramayah, T. (2020). Sampling weight adjustments in partial least squares structural equation modeling: Guidelines and illustrations. *Total Quality Management & Business Excellence*, 1–20.

Cheung, F., Tang, C. S. K., & Tang, S. (2011). Psychological capital as a moderator between emotional labor, burnout, and job satisfaction among school teachers in China. *International Journal of Stress Management*, *18*(4), 348.

Csikszentmihalyi, I. (1988). Flow in historical context: The case of the Jesuits. In M. Csikszentmihalyi & I. Csikszentmihalyi (Eds.), Optimal experience. *Psychological studies of flow in conciousness*. Cambridge University Press.

Curry, L. A., Snyder, C. R., Cook, D. L., Ruby, B. C., & Rehm, M. (1997). Role of hope in academic and sport achievement. *Journal of Personality and Social Psychology*, *73*(6), 1257.

Daniel, J. L. (2015). *Workplace spirituality and stress: Evidence from Mexico and US*. Management Research Review.

Dehler, G. E., & Welsh, M. A. 2003. The experience of work: Spirituality and the new workplace. In R. A. Giacalone and C. L. Jurkiewicz (Eds.), *Handbook of workplace spirituality and organizational performance*. M. E. Sharpe.

Dirkx, J. M. (2013). Leaning in and leaning back at the same time: Toward a spirituality of work-related learning. *Advances in Developing Human Resources*, *15*(4), 356–369.

Edwards, J. R., & Cooper, C. L. (1988). The impacts of positive psychological states on physical health: A review and theoretical framework. *Social Science & Medicine*, *27*(12), 1447–1459.

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4.

Fox, C., Webster, B. D., & Casper, W. C. (2018). Spirituality, psychological capital and employee performance: An empirical examination. *Journal of Managerial Issues*, 194–213.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218.

Fry, L. W. (2003). Toward a theory of spiritual leadership. *The Leadership Quarterly*, *14*(6), 693–727.

Garg, N., & Gera, S. (2019). Gratitude and leadership in higher education institutions. *Journal of Applied Research in Higher Education*, *12*(5), 915–926.

Garg, N., Punia, B. K., & Jain, A. (2019). Workplace spirituality and job satisfaction: Exploring mediating effect of organization citizenship behaviour. *Vision*, *23*(3), 287–296.

Garg, N., Gera, S., & Punia, B. K. (2021). Gratitude and turnover intention in e-commerce startups: Investigating the mediating role of organizational citizenship behaviour. *Journal of Electronic Commerce in Organizations (JECO)*, *19*(2), 34–54.

Giacalone, R. A., & Jurkiewicz, C. L. (2003). Right from wrong: The influence of spirituality on perceptions of unethical business activities. *Journal of Ethics*, *46*(1), 85–97.

Golparvar, M., & Jafari, M. (2013). Prediction of psychological capital through components of spirituality among nurses. *Iranian Journal of Psychiatric Nursing*, *1*(3), 35–44.

Hair, J. F.Jr, Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101–110.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513.

Imran, M., & Shahnawaz, M. G. (2020). PsyCap and performance: Wellbeing at work as a mediator. *Asia-Pacific Journal of Management Research and Innovation*, *16*(2), 93–102.

Jena, L. K., & Pradhan, R. K. (2015). Psychological capital and workplace spirituality: Role of emotional intelligence. *International Journal of Work Organisation and Emotion*, *7*(1), 1–15.

Karakas, F. (2010). Spirituality and performance in organizations: A literature review. *Journal of Business Ethics*, *94*(1), 89–106.

Kawalya, C., Munene, J. C., Ntayi, J., Kagaari, J., Mafabi, S., & Kasekende, F. (2019). Psychological capital and happiness at the workplace: The mediating role of flow experience. *Cogent Business & Management*, *6*(1), 1685060.

Kim, S., & Esquivel, G. B. (2011). Adolescent spirituality and resilience: Theory, research, and educational practices. *Psychology in the Schools*, *48*(7), 755–765.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration (IJEC)*, *11*(4), 1–10.

Kong, F., Tsai, C. H., Tsai, F. S., Huang, W., & De la Cruz, S. M. (2018). Psychological capital research: A meta-analysis and implications for management sustainability. *Sustainability*, *10*(10), 3457.

Krasikova, D. V., Green, S. G., & LeBreton, J. M. (2013). Destructive leadership: A theoretical review, integration, and future research agenda. *Journal of Management*, *39*(5), 1308–1338.

Lepisto, D. A., & Pratt, M. G. (2017). Meaningful work as realization and justification: Toward a dual conceptualization. *Organizational Psychology Review*, *7*(2), 99–121.

Lifeng, Z. (2007). Effects of psychological capital on employees' job performance, organizational commitment, and organizational citizenship behavior. *Acta Psychologica Sinica*, *2*, 18.

Lips-Wiersma, M., & Morris, L. (2009). Discriminating between 'meaningful work'and the 'management of meaning'. *Journal of Business Ethics*, *88*(3), 491–511.

Luthans, F. (2002). Positive organizational behavior: Developing and managing psychological strengths. *Academy of Management Perspectives*, *16*(1), 57–72.

Luthans, F., & Avolio, B. J. (2003). Authentic leadership development. *Positive Organizational Scholarship*, *241*, 258.

Luthans, F., Luthans, K. W., & Luthans, B. C. (2004). Positive psychological capital: Beyond human and social capital. *Business Horizons*, *47*(1), 45–50.

Luthans, F., Youssef, C. M., & Avolio, B. J. (2007). *Psychological capital: Developing the human competitive edge*. Oxford University Press.

Luthans, F., Avey, J. B., Avolio, B. J., Norman, S. M., & Combs, G. M. (2006). Psychological capital development: Toward a micro-intervention. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *27*(3), 387–393.

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common method bias in marketing: Causes, mechanisms, and procedural remedies. *Journal of Retailing*, *88*(4), 542–555.

Maddux, J. E. (Ed.). (2013). *Self-efficacy, adaptation, and adjustment: Theory, research, and application*. Springer Science & Business Media.

Mahipalan, M., & Sheena, S. (2019). Workplace spirituality and subjective happiness among high school teachers: Gratitude as a moderator. *Explore*, *15*(2), 107–114.

Manning, L. K. (2013). Navigating hardships in old age: Exploring the relationship between spirituality and resilience in later life. *Qualitative Health Research*, *23*(4), 568–575.

Masten, A. S., & Reed, M. G. J. (2002). Resilience in development. *Handbook of Positive Psychology*, *74*, 88.

McCormick, D. W. (1994). Spirituality and management. *Journal of Managerial Psychology*, 9(6), 5–8.

Michaelson, C., Pratt, M. G., Grant, A. M., & Dunn, C. P. (2014). Meaningful work: Connecting business ethics and organization studies. *Journal of Business Ethics*, *121*(1), 77–90.

Mitroff, I., & Denton, E. (1999). A study of spirituality in the workplace. *Sloan Management Review*, *40*(4), 83–92.

Nash, L. L., McLennan, S., & Blanchard, K. H. (2001). *Church on Sunday, work on Monday: The challenge of fusing Christian values with business life*. Jossey-Bass.

Pandey, A., & Gupta, R. K. (2008). Spirituality in management: A review of contemporary and traditional thoughts and agenda for research. *Global Business Review*, *9*(1), 65–83.

Mahoney, J., Gina M., & Graci, M. (1999). The meanings and correlates of spirituality: Suggestions from an exploratory survey of experts. *Death Studies*, *23*(6), 521–528.

Nash, M., & Stewart, B. (2005). Spirituality and hope in social work for social justice. *Currents: New Scholarship in the Human Services*, *4*(1).

Novitasari, D., Siswanto, E., Purwanto, A., & Fahmi, K. (2020). Authentic Leadership and Innovation: What is the Role of Psychological Capital? *International Journal of Social and Management Studies*, *1*(1), 1–21.

Pandey, A. (2017). Workplace spirituality: Themes, impact and research directions. *South Asian Journal of Human Resources Management*, *4*(2), 212–217.

Pandey, A., Gupta, V., & Gupta, R. K. (2019). Spirituality and innovative behaviour in teams: Examining the mediating role of team learning. *IIMB Management Review*, *31*(2), 116–126.

Pardasani, R., Sharma, R. R., & Bindlish, P. (2014). Facilitating workplace spirituality: Lessons from Indian spiritual traditions. *Journal of Management Development*, *33*(8/9), 847–859.

Pareek, U. (2002). *Training instruments for human resource development* (pp. 517–526). Tata McGraw-Hill.

Paul, M., & Saha, P. (2016). Spirituality as a potential psychological capital tool to deal with academic stress among management students. *IUP Journal of Organizational Behavior*, *15*(4), 46.

Pawar, B. S. (2009). Workplace spirituality facilitation: A comprehensive model. *Journal of Business Ethics*, *90*(3), 375–386.

Pawar, B. S. (2017). The relationship of individual spirituality and organizational spirituality with meaning and community at work: An empirical examination of the direct effects and moderating effect models. *Leadership & Organization Development Journal*, *38*(7), 986–1003.

Pedreira, L., & Mónico, L. D. S. M. (2013). Workaholism and psychological capital: Repercussions on workplace spirituality. *International Journal of Developmental and Educational Psychology*, *2*(1), 535–544.

Peres, J. F., Moreira-Almeida, A., Nasello, A. G., & Koenig, H. G. (2007). Spirituality and resilience in trauma victims. *Journal of Religion and Health*, *46*(3), 343–350.

Petchsawang, P., & McLean, G. N. (2017). Workplace spirituality, mindfulness meditation, and work engagement. *Journal of Management, Spirituality & Religion*, *14*(3), 216–244.

Peters, C., Amato, C. H., & Hollenbeck, C. R. (2007). An exploratory investigation of consumers' perceptions of wireless advertising. *Journal of Advertising*, *36*(4), 129–145.

Petersen, K., & Youssef-Morgan, C. M. (2018). The "left side" of authentic leadership: Contributions of climate and psychological capital. *Leadership & Organization Development Journal*, *39*(3), 436–452.

Pirkola, H., Rantakokko, P., & Suhonen, M. (2016). Workplace spirituality in health care: an integrated review of the literature. *Journal of Nursing Management*, *24*(7), 859–868.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539–569.

Powe, B. D. (1997). Cancer fatalism—Spiritual perspectives. *Journal of Religion and Health*, *36*(2), 135–144.

Pradhan, S., & Jena, L. K. (2017). Effect of abusive supervision on employee's intention to quit and the neutralizing role of meaningful work in Indian IT organizations. *International Journal of Organizational Analysis*, *25*(5), 825–838.

Rego, A., Sousa, F., Marques, C., & e Cunha, M. P. (2012). Authentic leadership promoting employees' psychological capital and creativity. *Journal of Business Research*, *65*(3), 429–437.

Reio Jr, T. G. (2010). The threat of common method variance bias to theory building. *Human Resource Development Review*, *9*(4), 405–411.

Rosso, B. D., Dekas, K. H., & Wrzesniewski, A. (2010). On the meaning of work: A theoretical integration and review. *Research in Organizational Behavior*, *30*, 91–127.

Rousseau, D. (2014). A systems model of spirituality: with William Simpson, "The Mystical Stance: The Experience of Self-Loss and Daniel Dennett's 'Center of Narrative Gravity'"; David Rousseau,"A Systems Model of Spirituality"; and David James Stewart,"The Emergence of Consciousness in Genesis 1–3: Jung's Depth Psychology and Theological Anthropology". *Zygon®*, *49*(2), 476–508.

Sarkar, A., & Garg, N. (2020). 'Peaceful workplace' only a myth? Examining the mediating role of psychological capital on spirituality and non-violence behaviour at the workplace. *International Journal of Conflict Management*, *31*(5), 709–728.

Sarkar, A., & Garg, N. (2021). A multilevel model exploring the relationships between individual spirituality, spiritual climate, and non-violent work behavior. *Conflict Resolution Quarterly*, *38*(3), 209–229.

Sarstedt, M., Hair Jr, J. F., Cheah, J. H., Becker, J. M., & Ringle, C. M. (2019). How to specify, estimate, and validate higher-order constructs in PLS-SEM. *Australasian Marketing Journal (AMJ)*, *27*(3), 197–211.

Sahoo, B. C., & Sia, S. K. (2015). Psychological capital and organisational commitment: Nature, structure and relationship in an Indian sample. *Asia-Pacific Journal of Management Research and Innovation*, *11*(3), 230–244.

Saxena, A., Garg, N., Punia, B. K., & Prasad, A. (2020). Exploring role of Indian workplace spirituality in stress management: A study of oil and gas industry. *Journal of Organizational Change Management*, *33*(5), 779–803.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, *4*(3), 219.

Schwartz, B. (1967). The social psychology of the gift. *American Journal of Sociology*, *73*, 1–11.

Seligman, M. E. (1998). Building human strength: Psychology's forgotten mission. *APA Monitor*, *29*, 2.

Seligman, M. E. (2004). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment*. Simon and Schuster.

Sepahvand, T. (2019). The role of the workplace spirituality and resilience in prediction of psychological well-being of employee. *Journal of Psychology*, *23*(3), 320–334.

Sharma, A., Dwivedi, Y. K., Arya, V., & Siddiqui, M. Q. (2021). Does SMS advertising still have relevance to increase consumer purchase intention? A hybrid PLS-SEM-neural network modelling approach. *Computers in Human Behavior*, 106919.

Shinde, U., Nelson, H. J., & Shinde, J. (2018). To be or not to be: A multidimensional spirituality in the workplace. *Journal of Human Values*, *24*(3), 185–207.

Shiva, A., Narula, S., & Shahi, S. K. (2020). What drives retail investors' investment decisions? evidence from no mobile phone phobia (nomophobia) and investor fear of missing out (I–FOMO). *Journal of Content, Community and Communication*, *10*(6), 2–20.

Simmons, B. L., & Nelson, D. L. (2001). Eustress at work: The relationship between hope and health in hospital nurses. *Health Care Management Review*, *26*(4), 7–18.

Singh, K., Junnarkar, M., & Kaur, J. (2016). *Measures of positive psychology*. Development and Validation, Springer.

Snyder, C. R. (2002). Hope theory: Rainbows in the mind. *Psychological Inquiry*, *13*(4), 249–275.

Snyder, C. R., Irving, L., &Anderson, J. R. (1991). Hope and health: Measuring the will and the ways. In C. R. Snyder & D. R. Forsyth (Eds.) *Handbook of social and clinical psychology: The health perspective* (pp. 285–305). Pergamon Press.

Soni, K., & Rastogi, R. (2019). Psychological capital augments employee engagement. *Psychological Studies*, *64*(4), 465–473.

Stajkovic, A. (2003). *Introducing positive psychology to work motivation: Development of a core confidence model*. Paper presented at the Academy of Management National Meeting, Seattle, WA.

Suárez, S. A. (2015). *Effects of spirituality in the workplace*. In Allied Academies International Conference. Academy of Organizational Culture, Communications and Conflict. Proceedings (Vol. *20*, No. 1, p. 9). Jordan Whitney Enterprises.

Thien, L. M. (2020). Assessing a second-order quality of school life construct using partial least squares structural equation modelling approach. *International Journal of Research & Method in Education*, *43*(3), 243–256.

Van Cappellen, P., Edwards, M. E., & Fredrickson, B. L. (2021). Upward spirals of positive emotions and religious behaviors. *Current Opinion in Psychology*, *40*, 92–98.

Vasconcelos, A. F. (2018). Workplace spirituality: Empirical evidence revisited. *Management Research Review*, *41*(7), 789–821

Vilariño del Castillo, D., & Lopez-Zafra, E. (2022). Antecedents of psychological capital at work: A systematic review of moderator–mediator Effects and a new integrative proposal. *European Management Review*, *19*(1), 154–169.

Weber, T. (2001). Gandhian philosophy, conflict resolution theory and practical approaches to negotiation. *Journal of Peace Research*, *38*(4), 493–513.

Youssef, C. M., & Luthans, F. (2007). Positive organizational behavior in the workplace: The impact of hope, optimism, and resilience. *Journal of Management*, *33*(5), 774–800.

Youssef, C. M., & Luthans, F. (2012). Psychological capital: Meaning, findings and future directions. In K. S. Cameron & G. M. Spreitzer (Eds.), *The Oxford handbook of positive organizational scholarship* (pp. 17–27). Oxford University Press.

Youssef-Morgan, C. M., & Luthans, F. (2015). Psychological capital and wellbeing. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *31*(3), 180–188.

Yu, X., Li, D., Tsai, C. H., & Wang, C. (2019). *The role of psychological capital in employee creativity*. Career Development International.

Zohar, D., & Marshall, I. (2000). *SQ: Spiritual intelligence, the ultimate intelligence*. Bloomsbury Publishing.