SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(54th Edition)

JUNE 2017

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 **Fifty fourth** 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 **June 2017**. 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 comparative analysis of operational performance of Cellular Mobile Telephone Service Providers in the Delhi working area using an approach of fuzzy ELECTRE, ***3.Pravin Kumar,*** Rajesh K. Singh and ***8.Karishma Kharab***, Mechanical, Production and Industrial, DTU

2. A Simple and Efficient Method for Removal of Phenolic Contaminants in Wastewater Using Covalent Immobilized Horseradish Peroxidase, Rajesh Ahirwar, ***3.Jai G. Sharma***, Bhanumati Singh, Krishan Kumar, Pradip Nahar and ***6.Saroj Kumar***, Biotechnology, DTU

3. An Analysis of Selected Parameters in an Energy Efficient Sustainable Manufacturing System Using AHP, Subrata Kumar Patra, Tilak Raj and ***3.B.B. Arora***, Mechanical, DTU

4. Analytical modeling to study the effect of hydrogen plasma on the growth of multi-walled carbon nanotubes, ***7.U Sharmaa***, ***7.N Gupta***, ***7.R Gupta*** and ***3.S C Sharma***, Applied Physics, DTU

5. Anticancer property of green material through computational approach, ***6.Nupur Jauhari***, Gajula MNV Prasad, Neelam Sharma and ***3.Navneeta Bharadvaja***, Biotechnology, DTU

6. Availability analysis of mechanical systems with condition-based maintenance using semi-Markov and evaluation of optimal condition monitoring interval, ***3.Girish Kumar***, Vipul Jain and O. P. Gandhi, Mechanical, DTU

7. Color texture recognition by color information fusion using the non-extensive entropy, ***3.Seba Susan*** and Madasu Hanmandlu, Information Technology, DTU

8. Condenser Heat Recovery For Combined Cooling-Heating and Power Generation using Isentropic Fluid, Kaushalendra K. Dubey and ***3.R. S. Mishra***, Mechanical, DTU

9. DEGRADATION OF PHORBOL ESTERS IN JATROPHA SEED CAKE BY PSEUDOMONAS AERUGINOSA DS1, SHILPI AHLUWALIA[1], ***3.JAI GOPAL SHARMA[1]*** and ***3.PUSHPENDRA SINGH[2]***, Biotechnology[1] and Mechanical[2], DTU

10. Determination of hierarchical relationships among sustainable development goals using interpretive structural modeling, ***3.Pravin Kumar***, Faisal Ahmed, Rajesh Kumar Singh and Prerna Sinha, Mechanical Engineering, DTU

11. Development of a paper-based electrochemical immunosensor using an antibody-single walled carbon nanotubes bio-conjugate modified electrode for label-free detection of foodborne pathogens, Jyoti Bhardwaj, Sivaranjani Devarakonda, ***6.Suveen Kumar*** and Jaesung Jang, Biotechnology, DTU

12. Effect of temperature on digestive physiology, immune-modulatory parameters, and expression level of Hsp and LDH genes in Catla catla (Hamilton, 1822), ***3.JaiGopal Sharma***, Samar Pal Singh and Rina Chakrabarti, Biotechnology, DTU

13. Electronic band-offsets across Cu2O/BaZrO3 heterojunction and its stable photo-electro-chemical response: First-principles theoretical analysis and experimental optimization, Dipika Sharma, ***7.Rishibrind Kumar Upadhyay***, Biswarup Satpati, Vibha R. Satsangi, Rohit Shrivastav, Umesh V. Waghmare and Sahab Dass, Applied Physics, DTU

14. Energy Efficient Probabilistic Clustering Technique for Data Aggregation in Wireless Sensor Network, ***3.Rajesh K, Yadav***, ***3.Daya Gupta*** and D. K. Lobiyal, Computer Engineering, DTU

15. Face Detection and Extraction from Low Resolution Surveillance Video Using Motion Segmentation, Vikram Mutneja and ***7.1.Dr. Satvir Singh***, Electronics & Communication, DTU

16. Grey relational analysis coupled with principal component analysis for optimization of the cyclic parameters of solar driven organic rankine cycle,

Deepak Tiwari, Ahmad Faizan Sherwani, MOHAMMAD ASJAD and ***3.Akhilesh Arora***, Mechanical, DTU

17. Group Authorization using Threshold Signatures for Medical Procedures, ***6.Preeti Nagar*** and ***3.Divyashikha Sethia***, Computer, DTU

18. Holographic Dark Energy Model in Brans-Dicke Theory with Future Event Horizon, ***6.Milan Srivastava*** and ***3.C. P. Singh***, Applied Mathematics, DTU

19. Integrated scheduling of part and tool in a flexible manufacturing system using modified genetic algorithm, Naveen Kumar, Pankaj Chandna and ***3.Dheeraj Joshi,*** DTU

20. Intelligent selection of threshold in covariance-based spectrum sensing for cognitive radio networks, ***7.1.Chhagan Charan*** and Rajoo Pandey, ECE, DTU

21. International journal of advanced production and industrial engineering, ***8.Sushil Yadav, 8.Lakshay, 8.Chitransh Joshi, 8.Mandeep Singh, 6.Gurmeet Singh Bhatia , 3.Ranganath M. Singari, 3.Nand Kumar,*** production and industrial, DTU

22. Lifetime quenching due to surface wake field, ***6.Gaurav Sharma***, ***3.Nitin K. Puri*** and T. Nandi, Applied Physics, DTU

23. Microencapsulation of reactive amine by interfacially engineered epoxy microcapsules for smart applications, ***7.Manorama Tripathi***, ***6.Devendra Kumar*** and · Prasun Kumar Roy, Applied Chemistry, DTU

24. Optimized MapFile based Storage of Small files in Hadoop, ***6.Divyashikha Sethia*** and ***7.Shalini Sheoran***, Computer, DTU

25. Performance Analyses and Optimisation of Energy Recovery Systems for Marine Applications, O P Sharma, S C Kaushik and ***3.K Manjunath,*** Mechanical, DTU

26. Quantitative evaluation of web metrics for automatic genre classification of web pages, ***3.Ruchika Malhotra*** and Anjali Sharma, Computer Science, DTU

27. Spectroscopic studies of Sm3þ ions activated lithium lead alumino borate glasses for visible luminescent device applications, ***6.Nisha Deopa*** and ***3.A.S. Rao***, Applied Physics, DTU

28. Terahertz pulses in a periodic plasma channel via modulated electron beam, ***6.Pratibha Malika***, ***3.Suresh C Sharma*** and ***3.Rinku Sharma***, Applied Physics, DTU

29. Terahertz radiation generation by using modulated electron beam with plasma wave wiggler, ***6.J Panwara***, ***3.S C Sharma*** and ***3.R Sharma***, Applied Physics, DTU

30. The versatility of an entropy inequality for the robust computation of convection dominated problems, Balaji Srinivasan and ***3.Vivek Kumar***, App Maths, DTU

31. Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique, ***6.Gaurav Sharma***, ***3.Nitin K. Puri*** and T. Nandi, Applied Physics, DTU

32. Theoretical investigation on the effect of nitrogen doping on the growth and field emission properties of the plasma-grown graphene sheet, ***7.N Gupta***, ***3.S C Sharma*** and ***7.R Gupta***, Applied Physics, DTU

33. Theoretical study to investigate the impact of plasma parameters on the catalyst nanoparticle growth, ***7.R Gupta***, ***3.S C Sharma*** and ***7.N Gupta***, App Physics, DTU

34. Thermal modelling of three stage vapour compression cascade refrigeration system using entropy generation principle for reducing global warming and ozone depletion using ecofriendly refrigerants for semen preservation, ***3.R S Mishra***, Mechanical, Production & Industrial and Automobiles, DTU

35. Thermodynamic (Energy-exergy) Analysis of Hydrogen Liquefaction System, ***6.Devender Kumar*** and ***3.R.S Mishra***, Mechanical, DTU

36. Thermodynamic (Exergy-energy) Analysis of a Low Pressure Improved Claude Cycle Using Control Volume Technique For Liquefaction of Gases, ***6.Devender Kumar*** and ***3.R.S Mishra***, Mechanical, DTU

37. Thermodynamic analysis and optimization of a supercritical CO2 regenerative recompression Brayton cycle coupled with a marine gas turbine for shipboard waste heat recovery, O P Sharma, S C Kaushik and ***3.K Manjunath***, Mechanical, DTU

38. Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, ***7.Nikita Gupta*** and ***3.Rachana Garg***, Electrical, DTU

39. Use of Data Mining Techniques in Assessing Student and Faculty Needs, Archna Bajaj and ***3.Dr. S. K. Saxena,*** Computer Science, DTU

1. *Chancellor*

2. *Pro Vice Chancellor* ***2.1.*** *Ex Pro Vice Chancellor*

3. *Faculty* ***3.1.*** *Ex Faculty*

4. *Teaching-cum-Research Fellow* ***4.1.****Asst. Librarian*

5. *Alumni*

6. *Research Scholar* ***6.1.*** *Ex Research Scholar*

7. *PG Scholar* ***7.1.*** *Ex PG Scholar*

8. *Undergraduate Student* ***8.1.*** *Ex Undergraduate Student*

Applied Soft Computing 59 (2017) 438–447

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A comparative analysis of operational performance of Cellular Mobile Telephone Service Providers in the Delhi working area using an approach of fuzzy ELECTRE

Pravin Kumar[a,*], Rajesh K. Singh[b], Karishma Kharab[c]

[a] Department of Mechanical Engineering, Delhi Technological University, Delhi, 110042, India
[b] Management Development Institute, Gurgaon, Haryana, 122007, India
[c] Department of Production and Industrial Engineering, Delhi Technological University, Delhi, 110042, India

ARTICLE INFO

Article history:
Received 22 November 2015
Received in revised form 24 April 2017
Accepted 10 June 2017

Keywords:
Fuzzy logic
ELECTRE
Cellular mobile telephone service provider
Performance measurement
Outranking

ABSTRACT

The purpose of this research paper is to develop a framework to analyze the operational performance of cellular mobile telephone service providers (CMTSP) in the Delhi working area, India. Delhi area has the highest teledensity in India. The present study uses a fuzzy ELECTRE (Elimination and Choice Expressing Reality) approach to compare the performance of cellular mobile telephone service providers. The data for the analysis have been taken from the Telecom Regulatory Authority of India (TRAI), April–June 2015. Most of the data vary from one month to another. Therefore, these small ranges of variation in the data are incorporated in this study using the fuzzy number. Total six major telecom service providers are considered in this analysis. The findings of the study suggest that the performances of Airtel and Reliance communications are in the first rank; Vodafone and Idea are in the second rank; Aircel is in the third rank, and Mahanagar Telephone Nigam Limited (MTNL) is outranked by all the other service providers. The performance is analyzed on the basis of three major parameters consisting six criteria, i.e. network availability (Base Transceiver Station accumulated downtime), connection accessibility (call setup success rate, channel congestion, and traffic channel congestion), and connection retainability (call drop rate and connection with good voice quality). It has been observed that MTNL is outranked due to the highest traffic channel congestion and highest call drop rate. Using this framework, a decision maker can develop a strategy to improve the performance by benchmarking operational parameters. The research analysis is limited to only GSM (Global System for Mobile communication) service providers in the Delhi working area, including Ghaziabad, Noida, Faridabad, and Gurgaon.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Telecommunication has become one of the fast-growing industries over the last two decades. The evolution of telecommunication sectors has affected growth at all levels, from individual firms to organizations at the regional and national level. The inventions of the mobile phone and the internet have given the impetus to dramatic economic growth [1]. Telecommunication is one of the prime support services needed for rapid growth and modernization of the business economy [2]. There has been a remarkable change in the world economy because of a constant development and up gradation of information and communication technologies. In particular, the telecommunication sector has triggered a revolution in shaping up the process of global changes to capture the opportunities. Furthermore, the growth in the telecommunication sector has also caused a huge increase in the amount of cross-border information flows, thereby reducing transaction costs and stimulating consumer demand for world-class products, services, and brands [3,4].

The telecommunication sector in India has been undergoing a massive expansion for over two decades in the breadth and choice of services and depth of penetration in urban, semi-urban and rural areas [5]. Delhi working area has the highest teledensity equal to 237.02 whereas the overall teledensity of India is 79.56. Total wireless telephone subscribers in India are 973.35 million. India is currently the second-largest telecommunication market in the world. India's telephone subscriber base expanded at a CAGR (Cumulative annual growth rate) of 19.22 percent to 1002 million over FY07–15. Teledensity (the number of telephone connections for every hundred individuals) increased from 18.3 in FY2007 to

* Corresponding author.
E-mail address: pravin.dce@gmail.com (P. Kumar).

http://dx.doi.org/10.1016/j.asoc.2017.06.019
1568-4946/© 2017 Elsevier B.V. All rights reserved.

79.56 in FY 2015. In May 2015, total telephone subscription stood at 1002 million. Urban regions account for 55.76 per cent of telecom subscriptions, while rural areas constitute the remaining [6].

As per TRAI [7], the telecommunication sector witnessed substantial growth in the number of subscribers during the year 2013-14. At the end of the financial year, the subscriber base was 933.00 million, out of which 904.51 million were wireless subscribers. During the year, wireless subscriber base recorded an increase of 36.71 million, while the overall teledensity increased from 73.32 to 75.23. The year also saw an increase in rural teledensity from 41.02 to 43.96 while the urban teledensity decreased to 145.78 from 146.96. As per the data reported by the service providers, by the end of March 2014 about 117.01 million mobile subscribers have submitted their requests to different service providers for porting their mobile number.

Proactive policies of Government of India, such as opening up the sector to private players and competition, unbundling the policy, regulatory and operational roles of the government, removal of restrictions on foreign investments coupled with viewing reforms as a continuous process have created an environment conducive to growth. Although, these reforms enabled the induction of upgraded technologies in this sector, but it still lacks comprehensive performance frameworks suitable for the current market scenario of dynamic customer needs. Therefore, the performance framework is urgently needed for rigorous and in-depth comparative analysis of different telecom operators in India. This kind of framework will also help in creating a competitive environment for the sustainable growth of telecom sector. In India, a large number of telecom service providers are working with a variety of services. The Delhi working area being the highest teledensity zone has been chosen as the scope of this research. To minimize the diversity of the services, only GSM cellular mobile telephone service providers in Delhi working area are considered for the performance analysis. Major objectives of this study are:

- To identify parameters for measuring the performance of telecom sector.
- To propose and illustrate a framework for comparative analysis of telecom service providers.

The rests of the paper have been organized in the following way: the second section reviews the Indian mobile sector and also discusses the teledensity and the rapid growth over the period in India. The third section provides the detailed methodology of fuzzy ELECTRE used for ranking of the telecom operators. The fourth section discusses the performance parameters and variables affecting the consumer's preference and performance analysis of the case companies. The fifth section concludes the research work.

2. Literature review

Due to the fast growth rate of the telecommunication industry, it has been in the focal area of the research for some researchers and practitioners in the recent past. Kumar & Kumar [5] used Analytic Hierarchy Process and Technique for Order of Preference by Similarity to Ideal Solution (AHP-TOPSIS) integrated tool to assess the performance of telecom service providers (TSPs) working in Delhi metropolis. They have used the quality of service (QOS) parameters for both the cellular and wire-line connections. In the present paper, we have used the same QOS parameters related to a cellular connection only. Kumar et al. [8] used a fuzzy Analytic Hierarchy Process (FAHP) and Data Envelope Analysis (DEA) integrated model for the assessment of the relative efficiency of Telecom Service providers. They have used BTS downtime, call success rate (CSR), call drop rate (CDR), and Good voice quality (GVQ) as input parameters and Number of subscribers, billing credibility, complaint resolved within four weeks, and percentage of call answered in customer care as output parameters in DEA. They incorporated the variable affecting consumers' preferences and perceptions. Saxena et al. [9] used benchmarking as a method of performance evaluation that assesses the potential efficiency improvements within telecom service providers.

Debnath and Shankar [4] used Data Envelopment Analysis (DEA) to compare the relative efficiency of mobile service providers in a few states of India and identified the frontiers and the poor service providers which could be very useful in improving their efficiency and performance. Nigam et al. [2] benchmarked the performance of Indian mobile telecommunication service providers for measuring the relative efficiencies based on DEA. They highlighted the inefficient utilities of telecom service providers based on the analysis that can be developed strategic plans to improve the performance. Tsai et al. [10] reconciled the diverse efficiency measures to characterize the productivity in terms of efficiency of 39 Forbes 2000 ranked leading global telecom operators by applying the DEA and observed that the telecom operators having top-rank in Forbes are not same found through CCR efficiency measures. Giokas and Pentzaropoulos [11] used data envelopment analysis for measuring the efficiency of telecom companies in Greece and investigated the efficient and inefficient companies. Zhu [12] used DEA for evaluating the relative efficiency in Korean telecom industries with strong ordinal input/output. Yang and Chang [13] used the Data Envelopment Analysis to calculate the efficiency of Taiwan telecom companies after pre and post-privatization. Resende [14] used the DEA model for efficiency measurement in the US telecom market and compared the result with stochastic frontier analysis for obtaining moderate consistency in these approaches for making a new regulation in telecom industries.

Joshi [15] worked to improve the revenue generation for the company for which he used factor analysis to group the attributes of the customer experience into common factors that affect the customer experience for cellular mobile services and concluded that customer retention can be achieved by identifying maximum revenue generating customers and managing the customer experience for such profitable customers. Majumdar [16] showed that encouraging the entry of competitors had led quite substantially to the enhancement of efficiencies for the incumbent firms in a two-sided network context of telecommunication firms in the USA using panel data regression model.

Yadav [17] developed a total interpretive structural modeling (TISM) of strategic factors related to performance management in the context of Indian Telecom Service Providers taking dual perspectives into account: enterprise perspective and subscribers' perspective using flexible strategy game-card as the basis to develop the strategic performance management model. The strategic factors have been identified through thematic content analysis and the hierarchical structure of factors has been developed using TISM and concluded that the hierarchical structure portrays the leading and lagging factors of performance.

Many researchers have already used the DEA as the efficiency analysis tools in the telecom industry. But, very few studies have been identified in the past research that are based on the outranking method to rank the performance of the telecom operators considering the operational quality parameters such as network availability (BTS accumulated downtime), connection accessibility (call setup success rate, channel congestion, and TCH congestion), and connection retainability (call drop rate and connection with good voice quality). These parameters have also been used as inputs for the efficiency analysis by Kumar et al. [8]. ELECTRE has already been in use as a decision-making tool in the area of engineering and management. Bottero et al. [18] used ELECTRE to deal multiple criteria environmental problem with interaction effects between

criteria. Zak and Kruszynski [19] applied analytic Hierarchy Process (AHP) and ELECTRE III/IV for the evaluation of urban transportation project. They considered it as a multi-criteria problem. Ka [20] used FAHP and ELECTRE for selection of China dry port location. Similarly, Montazer et al. [21] used fuzzy ELECTRE III for vendor selection. Tolga [22] used fuzzy ELECTRE for software development projects [23].

In this paper, authors have used fuzzy ELECTRE to rank the cellular mobile telephone service providers in the Delhi working area. The Key contribution of the paper is to propose an approach for the ranking of various cellular mobile telecom service providers in the Delhi working area based on operational performance. Other multi-criteria decision-making tools provide some ranks to all the service providers. This will help in outranking the poor performers without providing any rank. Therefore, it may help consumers to consider only the ranked service providers in Delhi working area. Additionally, it may help the TSPs (Telecom Service Providers) to improve the services by benchmarking the top rankers.

3. Research Methodology

Six TSPs have been considered in this study. For a ranking of these TSPs in terms of operational performance, Fuzzy ELECTRE approach has been used. Major steps of this approach have been discussed as follows:

3.1. Fuzzy ELECTRE

Different variants of ELECTRE method have been in use for the decision-making process. In this paper, ELECTERE is integrated with fuzzy sets to incorporate the small variations in performances [24].

Step 1: Preparation of fuzzy decision matrix

To prepare a fuzzy decision matrix of 'm' alternatives and 'n' attributes, at the first fuzzy importance of the attributes is determined. The fuzzy importance of the attributes or the operational parameters are determined on the basis of consumers opinion. The operational performance of the telecom companies in the months of April, May, June 2015 are taken from the TRAI report [6]. All these three values are arranged in sequence as $x_{ij}^{l} \le x_{ij}^{\gamma} \le x_{ij}^{u}$. Where $i = 1,2,\ldots m$ is the number of alternatives, $j = 1,2,\ldots n$ is the number of attributes, and $\tilde{x}_{ij}$ is the rating of i^{th} alternative against the j^{th} attributes. Here l indicates the lowest value of x_{ij}, y indicates the middle value of x_{ij}, and u indicates the largest value of x_{ij} among all the three values of x_{ij} for the months of April, May, and June2015. The data is taken from TRAI for the second quarter of 2015 (April-June) and minor variations in the data during these three months have been observed. To incorporate these variations for realistic analysis, fuzzy sets are used with ELECTRE methodology. These three values are considered as a triangular fuzzy number because the variations are not much significant and it has been assumed that the middle value is the best value having membership function's value equal to 1. The decision matrix can be represented as:

$$\tilde{U} = \begin{array}{c} \begin{matrix} \tilde{W}_1 & \tilde{W}_2 & \cdots & \tilde{W}_n \end{matrix} \\ \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \cdots & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & \cdots & \tilde{x}_{2n} \\ . & . & \cdots & . \\ \tilde{x}_{m1} & \tilde{x}_{m2} & \cdots & \tilde{x}_{mn} \end{bmatrix} \end{array} \quad (1)$$

Where the importance rating of the criteria is calculated as:

$\tilde{W}_j = \frac{1}{k}\sum \tilde{W}_{jk}$; Here, k is the number of consumers.

The consumer uses the five linguistic variables for the rating of importance of the criteria as Least Important, Less Important, Important, Very Important, and Most Important. These variables are converted into the quantitative scale as shown in Fig. 1.

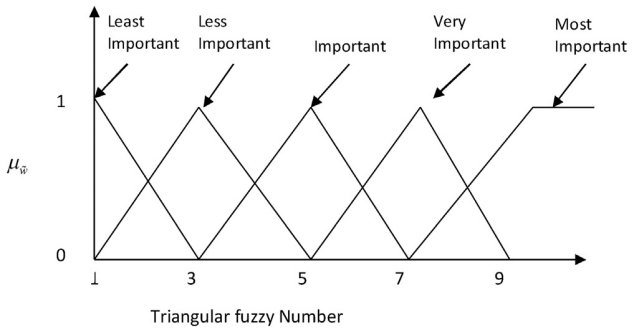


Fig. 1. Linguistic Scale for importance rating of criteria.

Step 2: preparation of normalized decision matrix

Linear scale normalization is applied to ensure that all values as the elements of the decision matrix have homogeneous and comparable units. Moreover, this transformation guarantees that every triangular fuzzy number belongs to [0,1] which reduces the complexity in the mathematical operations. The normalized fuzzy decision matrix is shown in Eq. (2).

$$\tilde{R} = \left[\tilde{r}_{ij}\right]_{m \times n}, \quad (2)$$

where

$$\tilde{r}_{ij} = \left(r_{ij}^{l}, r_{ij}^{\gamma}, r_{ij}^{u}\right) = \left(\frac{x_{ij}^{l}}{c_j^*}, \frac{x_{ij}^{\gamma}}{c_j^*}, \frac{x_{ij}^{u}}{c_j^*}\right), \quad (2a)$$

$$c_j^* = \max_i (x_{ij}^{u}),\ j \in B, \text{ and}$$

B is the set of benefit attributes, i = 1,2. . .,m and j = 1,2,. . .,n.

Eq. (2a) is used to normalize the ratings with respect to the criteria which belong to benefit attribute, and

$$\tilde{r}_{ij} = \left(r_{ij}^{l}, r_{ij}^{\gamma}, r_{ij}^{u}\right) = \left(\frac{a_j^-}{x_{ij}^{u}}, \frac{a_j^-}{x_{ij}^{\gamma}}, \frac{a_j^-}{x_{ij}^{l}}\right), \quad (2b)$$

$$a_j^- = \min_i (x_{ij}^{l}),\ j \in C,$$

C is the set of cost attributes, i = 1,2. . .,m and j = 1,2,. . .,n.

Eq. (2b) is used to normalize the ratings with respect to the criteria which belong to cost attribute.

Step 3: compute the weighted normalized matrix

Multiply the standard rating of the alternatives with the importance of corresponding attributes to compute the weighted, normalized matrix using Eq. (3).

$$\tilde{v}_{ij} = \left(\tilde{r}_{ij}\right) \times \left(\tilde{w}_j\right) = \left(r_{ij}^{l}, r_{ij}^{\gamma}, r_{ij}^{u}\right) \times \left(w_j^{l}, w_j^{\gamma}, w_j^{u}\right). \quad (3)$$

Step 4: calculate the distance between any two alternatives

Hamming distance between two fuzzy numbers $(\tilde{A},\tilde{B})$ is defined by Hamming [25]

$$\int_R |\mu_{\tilde{A}}(x) - \mu_{\tilde{B}}(x)| \quad (4)$$

where R is the set of real numbers.

But, for simplicity in calculation, the distance between two fuzzy numbers, especially in ELECTRE method can be represented by the Eq. (5) [26,27]

$$d\left(\tilde{A}, \tilde{B}\right) = \sqrt{\sum_{k=1}^{3}(A_k - B_k)^2}. \quad (5)$$

The concordance and discordance matrices are prepared from weighted normalized fuzzy decision matrix and pairwise comparison of the alternatives. Considering two alternatives A_x and A_y, the concordance set is formed as $J^C = \{J|\tilde{v}_{xj} \geq \tilde{v}_{yj}\}$where J^C is the concordance coalition of the attributes in which A_xSA_y and the discordance set is defined as$J^D = \{J|\tilde{v}_{xj} \leq \tilde{v}_{yj}\}$ where J^D is the discordance coalition and it is against the assertion A_xSA_y. Note that S is the outranking relation and A_xSA_y means that "A_x is at least as good as A_y."

In order to compare any two alternatives A_x and Ay, with respect to each attribute, and to define the concordance and discordance sets, we specify the least upper bound of the alternatives, max $\left(\tilde{v}_{xj}, \tilde{v}_{yj}\right)$ and then the distance is used with the following assumptions:

$$\tilde{v}_{xj} \geq \tilde{v}_{yj} \Leftrightarrow d(\max(\tilde{v}_{xj}, \tilde{v}_{yj}), \tilde{v}_{yj}) \geq d(\max(\tilde{v}_{xj}, \tilde{v}_{yj}), \tilde{v}_{xj}), \text{ and} \quad (6)$$

$$\tilde{v}_{xj} \leq \tilde{v}_{yj} \Leftrightarrow d(\max(\tilde{v}_{xj}, \tilde{v}_{yj}), \tilde{v}_{yj}) \leq d(\max(\tilde{v}_{xj}, \tilde{v}_{yj}), \tilde{v}_{xj}). \quad (7)$$

Step 5: Construct the concordance sets of criteria, and prepare the concordance matrix

The concordance set is the group of attributes for which$\tilde{v}_{xj} \geq \tilde{v}_{yj}$. The following concordance matrix is prepared in which the elements are the fuzzy summation of the fuzzy importance weights for all the attributes in the concordance set as shown in Eq. (8).

$$\tilde{C} = \begin{bmatrix} \tilde{c}_{11} & \cdots & \tilde{c}_{1x} & \cdots & \tilde{c}_{1(m-1)} & \tilde{c}_{1m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \tilde{c}_{x1} & \cdots & \cdots & \cdots & \tilde{c}_{x(m-1)} & \tilde{c}_{xm} \\ \cdots & \cdots & \cdots & \tilde{c}_{xy} & \cdots & \cdots \\ \tilde{c}_{(m-1)1} & \cdots & \tilde{c}_{(m-1)y} & \cdots & \cdots & \cdots \\ \tilde{c}_{m1} & \cdots & \cdots & \cdots & \tilde{c}_{m(m-1)} & \tilde{c}_{mm} \end{bmatrix} \quad (8)$$

$$\text{where } \tilde{c}_{xy} = \left(c_{xy}^{l}, c_{xy}^{\gamma}, c_{xy}^{u}\right) = \left(\sum_{J \in J^c} w_J^{l}, \sum_{J \in J^c} w_J^{\gamma}, \sum_{J \in J^c} w_J^{u}\right).$$

Now, determine the concordance level as:

$$\bar{\tilde{c}} = \sum_{x=1}^{m} \sum_{y=1}^{(m-1)} \frac{\tilde{c}_{xy}}{m(m-1)}. \quad (9)$$

Step 6: prepare a Boolean matrix E from the concordance matrix

$$E = \begin{bmatrix} e_{11} & \cdots & e_{1x} & \cdots & e_{1(m-1)} & e_{1m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ e_{x1} & \cdots & \cdots & \cdots & e_{x(m-1)} & e_{xm} \\ \cdots & \cdots & \cdots & e_{xy} & \cdots & \cdots \\ e_{(m-1)1} & \cdots & e_{(m-1)y} & \cdots & \cdots & \cdots \\ e_{m1} & \cdots & \cdots & \cdots & e_{m(m-1)} & e_{mm} \end{bmatrix} \quad (10)$$

Where

$$\begin{cases} \tilde{c}_{xy} \geq \bar{\tilde{c}} \Leftrightarrow e_{xy} = 1, \text{ and} \\ \tilde{c}_{xy} \leq \bar{\tilde{c}} \Leftrightarrow e_{xy} = 0. \end{cases}$$

Step 7: construct the discordance sets of criteria, and prepare the discordance matrix

The discordance set is the set of the attributes for which$\tilde{v}_{xj} \leq \tilde{v}_{yj}$. The following discordance matrix is prepared in which the elements are the fuzzy summation of the fuzzy importance weights for all the attributes in the discordance set as shown in Eq. (11).

$$D = \begin{bmatrix} d_{11} & \cdots & d_{1x} & \cdots & d_{1(m-1)} & d_{1m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ d_{x1} & \cdots & \cdots & \cdots & d_{x(m-1)} & d_{xm} \\ \cdots & \cdots & \cdots & d_{xy} & \cdots & \cdots \\ d_{(m-1)1} & \cdots & d_{(m-1)y} & \cdots & \cdots & \cdots \\ d_{m1} & \cdots & \cdots & \cdots & d_{m(m-1)} & d_{mm} \end{bmatrix} \quad (11)$$

Where

$$d_{xy} = \frac{\max\limits_{j \in J^D} |\tilde{v}_{xj} - \tilde{v}_{yj}|}{\max\limits_{j} |\tilde{v}_{xj} - \tilde{v}_{yj}|} = \frac{\max\limits_{j \in J^D} |d(\max(\tilde{v}_{xj}, \tilde{v}_{yj}), \tilde{v}_{xj}|}{\max\limits_{j} |d(\max(\tilde{v}_{xj}, \tilde{v}_{yj}), \tilde{v}_{yj})|}.$$

Now, determine the discordance level as:

$$\bar{d} = \sum_{x=1}^{m} \sum_{y=1}^{(m-1)} \frac{d_{xy}}{m(m-1)}. \quad (12)$$

Step 8: prepare a Boolean matrix F from the discordance matrix

$$F = \begin{bmatrix} f_{11} & \cdots & f_{1x} & \cdots & f_{1(m-1)} & f_{1m} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ f_{x1} & \cdots & \cdots & \cdots & f_{x(m-1)} & f_{xm} \\ \cdots & \cdots & \cdots & f_{xy} & \cdots & \cdots \\ f_{(m-1)1} & \cdots & f_{(m-1)y} & \cdots & \cdots & \cdots \\ f_{m1} & \cdots & \cdots & \cdots & f_{m(m-1)} & f_{mm} \end{bmatrix} \quad (13)$$

where

$$\begin{cases} d_{xy} \geq \bar{d} \Leftrightarrow f_{xy} = 0, \text{ and} \\ d_{xy} \leq \bar{d} \Leftrightarrow f_{xy} = 1. \end{cases}$$

Step 9: prepare the global matrix

The global matrix G is prepared by peer-to-peer multiplication of the elements of the matrices E and F, It can be represented as:

$$G = E \otimes F \quad (14)$$

Step 10: construct a decision graph and rank the alternatives

A decision graph is drawn using the global matrix G in order to determine the ranking order of the alternatives. Let A_x and A_y be two alternatives. There is an arc between the two alternatives from A_x to A_y, if alternative A_x outranks A_y, there is no arc between the two alternatives if alternatives A_x and A_y are incomparable, and there are two arcs between the two alternatives in both directions if these alternatives are indifferent.

4. Case illustration

In the Delhi working area, Aircel, Airtel, MTNL, Reliance Communications, and Vodafone are the major cellular mobile GSM telephone operators. These companies are considered in the analysis of operational performance. The operational performance parameters have been taken from the TRAI report. These parameters are shown in Fig. 2.

4.1. BTS (Base Transceiver Station) downtime

It is an indicator of network availability. Less downtime means more availability. A BTS consists of the antennas that relay radio messages, transceivers, duplexers, and amplifiers. A network may

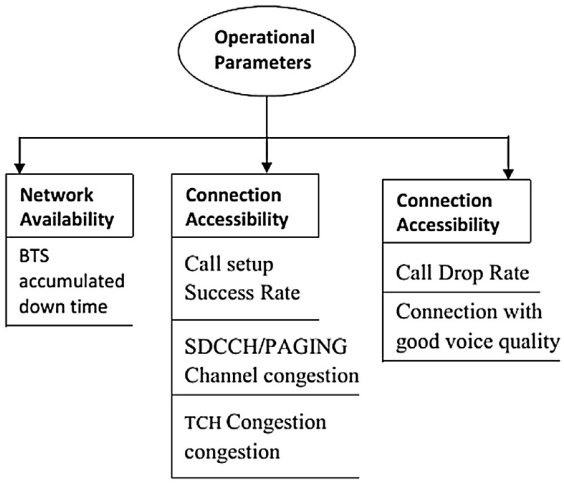


Fig. 2. Operational parameters for the performance measurement of CMTSP.

be any wireless technology, like Code Division Multiple Access (CDMA), Global System for Mobile Communications (GSM), Worldwide Interoperability for Microwave Access (WiMAX) or Wi-Fi. However, here only GSM is considered.

4.2. The call setup success rate (CSSR)

The call setup success rate (CSSR) is the fraction of the attempts to make a call that results in a connection to the dialed number (due to various reasons not all call attempts end with a connection to the dialed number). This fraction is usually measured as a percentage of all call attempts made.

4.3. SDC channel (SDCCH) congestion

The SDCCH stands for Stand-alone Dedicated Control Channel. *SDCCH Congestion* is defined as the probability of failure of accessing a stand-alone dedicated control channel during call setup. The SDCCH congestion rate is one of the important counters that indicate accessibility in circuit service. This counter provides the ratio of failed SDCCH seizures due to the busy SDCCH to the total requests for the SDCCH.

4.4. TCH (Traffic channel) congestion

Traffic channel in GSM carries user speech or data as other technology. GSM has two types of traffic channels TCH/FS and TCH/HS. TCH/FS stands for the traffic channel at full rate speech and TCH/HS stands for the traffic channel at half rate speech. The TCH congestion rate is the proportion of the number of TCH assignment failures to the number of TCH seizure requests. If the TCH congestion rate is high, the network service quality deteriorates.

4.5. The dropped-call rate (DCR)

It is the fraction of the telephone calls which, due to technical reasons, were cut off before the speaking parties had finished their conversational tone and before one of them had hung up (dropped calls) This fraction is usually measured as a percentage of all calls.

4.6. Connection with good voice quality

Sometimes the mobile network connects with unclear voice, i.e. problem in hearing the voice of the caller. Connection with good voice quality is a desirable operational parameter.

The values of the operational parameters of the cellular mobile telephone operators in the Delhi area are taken from the source TRAI (Aril-May 2015). The Delhi working area covers Delhi, Ghaziabad, Noida, Faridabad, Gurgaon. Table 1 shows the values of the operational parameters of the various cellular mobile telephone operators.

These operational parameters are considered as the criteria for the evaluation of the telephone service providers. All these parameters are important for performing the services, but variation in the performance of different operators (service providers) regarding these criteria forms a complexity in the ranking analysis. In this paper, the Delhi based customers' opinions are recorded for finding the importance of the criteria and five level variables in linguistic terms ("Least important", "Less important", "Important", "Very

Table 1
Operational parameters of Cellular Mobile Telephone Service Providers Delhi Metro Circle.

Name of Parameter	Benchmark	Audit Period	Aircel (A1)	Airtel (A2)	MTNL (A3)	Idea (A4)	RCom GSM (A5)	Vodafone (A6)
A. Network Availability								
1. BTS accumulated downtime	<=2%	Apr-2015	0.37%	0.01%	0.12%	0.01%	0.32%	0.12%
		May-2015	0.48%	0.01%	0.19%	0.01%	0.32%	0.15%
		June- 2015	0.42%	0.01%	0.19%	0.02%	0.39%	0.13%
B. Connection Accessibility								
2. Call setup Success Rate	>=95%	Apr- 2015	97.30%	99.85%	96.72%	99.80%	99.46%	99.43%
		May- 2015	97.54%	99.84%	97.25%	99.77%	99.47%	99.49%
		June- 2015	97.37%	99.83%	96.96%	99.76%	99.41%	99.38%
3. SDCCH/PAGING Channel congestion	<=1%	Apr- 2015	0.73%	0.03%	0.58%	0.66%	0.04%	0.04%
		May- 2015	0.82%	0.05%	0.57%	0.83%	0.04%	0.05%
		June- 2015	0.77%	0.06%	0.61%	0.69%	0.07%	0.03%
4. TCH Congestion	<=2%	Apr- 2015	1.60%	0.08%	1.75%	1.61%	0.10%	0.11%
		May- 2015	1.11%	0.06%	1.78%	1.71%	0.11%	0.10%
		June- 2015	1.59%	0.08%	1.81%	1.75%	0.12%	0.13%
C. Connection Retainability								
5. Call Drop Rate	<=2%	Apr- 2015	1.13%	0.48%	1.88%	0.60%	0.35%	1.06%
		May- 2015	1.07%	0.52%	1.80%	0.58%	0.32%	1.00%
		June- 2015	1.14%	0.57%	1/84%	0.59%	0.33%	1.06%
6. Connection with good voice quality	>=95%	Apr- 2015	97.19%	99.38%	96.94%	98.57%	98.56%	97.28%
		May- 2015	97.25%	99.39%	96.96%	98.64%	98.67%	97.38%
		June- 2015	97.25%	99.30%	96.96%	98.64%	98.65%	97.26%

(Source: Audit & Assessment of Qos For Qe-June 2015-Delhi Metro Circle http://trai.gov.in/WriteReadData/Qos/Document/201510160515381970748QoS%20Audit%20Report%20QE%20June%20%2015- Delhi%20Metro%20Circle. pdf).

Table 2
Scale for the importance rating of the criteria.

Importance on linguistic Scale	Fuzzy scale
Least important	(1,1,3)
Less important	(1,3,5)
Important	(3,5,7)
Very Important	(5,7,9)
Most important	(7,9,9)

Table 3
Common opinion given by customers for the importance of the criteria.

Criteria	Customers' common opinion	Fuzzy scale
C1 (BTS accumulated downtime)	Least important	(1,1,3)
C2 (Call setup success rate)	Most important	(7,9,9)
C3 (SDC channel congestion)	Important	(3,5,7)
C4 (TCH Congestion)	Important	(3,5,7)
C5 (Call drop rate)	Most important	(7,9,9)
C6 (Connections with good voice quality)	Very important	(5,7,9)

important", "Most important") are used to capture the vagueness in decision making as shown in Table 2 and also in Fig. 1

In Table 3, the levels of the linguistic variables shown are the common opinion of the customers (majority of opinions) and these are incidentally four levels of variables, however, customers were provided the questionnaire on five levels of linguistic variables as shown in Table 2. They had used all the five variables to rate the importance of the criteria, but Table 3 shows the most common response of the customers. Here, the majority of opinions is used to show the importance of the criteria to minimize the complexity of the calculations not using the average value of the fuzzy numbers. In this survey, more than 100 mobile phone users in Delhi were consulted to know the importance of the criteria.

Based on the problem domain, an appropriate linguistic variable set is chosen and used to describe the vague or imprecise knowledge. The number of levels of the linguistic variables determines the granularity of the uncertainty. According to the uncertainty degree that an expert qualifying a phenomenon has on it, the linguistic variable set chosen to provide his knowledge will have more or less levels. When different experts have different uncertainty degrees on the phenomenon, then several linguistic level sets with a different granularity of uncertainty are necessary [28].

It has been observed that in most of the research papers such as [29–35], 5-level of linguistic variables are frequently used for the importance rating. Therefore, this study has also considered five levels. However, the linguistic variables used in the decision making may be 3-levels, 5-levels, or 7-levels based on the nature and scope of the problem. With increasing the number of levels of linguistic variables, the result may move towards more accuracy, but at the same time, it may increase the difficulty for the decision makers in selection of exact variable among so many options. If the less number of levels of linguistic variables is used, then it may not capture the exact vagueness of states of the mind of the human beings in decision making and accuracy of the result may not be good. It may further impact the robustness of the model.

The ratings of the telephone service providers for the three months (April – June 2015) are considered as fuzzy numbers since it varies continuously. Variations in some parameters are very small i.e. negligible. Here, the fuzzy logic incorporates an even smaller range of variation very easily. Thus, fuzzy extended ELECTRE is used to outrank the poor performers and find the rank of the rest per-

Table 4
Fuzzy decision matrix of cellular mobile telecom service providers.

Criteria	Alternatives					
	A1 (Aircel)	**A2** (Airtel)	**A3** (MTNL)	**A4** (Idea)	**A5** (Rcom GSM)	**A6** (Vodafone)
C1	(0.37,0.42,0.48)	(0.01,0.01,0.01)	(0.12,0.19,0.19)	(0.01,0.01,0.02)	(0.32,0.32,0.39)	(0.12,0.13,0.15)
C2	(97.30,97.37,97.54)	(99.83,99.84,99.85)	(96.72,96.96,97.25)	(99.76,99.77,99.80)	(99.41,99.46,99.47)	(99.38,99.43,99.49)
C3	(0.73,0.74,0.82)	(0.03,0.05,0.05)	(0.57,0.58,0.61)	(0.63,0.69,0.83)	(0.04,0.04,0.07)	(0.03,0.04,0.05)
C4	(1.11,1.59,1.60)	(0.06,0.08,0.08)	(1.75,1.78,1.81)	(1.61,1.71,1.75)	(0.10,0.11,0.12)	(0.10,0.11,0.13)
C5	(1.07,1.13,1.14)	(0.48,0.52,0.57)	(1.80,1.84,1.88)	(0.58,0.59,0.60)	(0.32,0.33,0.35)	(1,1.06,1.06)
C6	(97.19,97.25,97.25)	(99.3,99.38,99.39)	(96.94,96.96,96.96)	(98.57,98.64,98.64)	(98.56,98.65,98.67)	(97.26,97.28,97.38)

Table 5
Normalized fuzzy decision matrix.

Criteria	Alternatives					
	A1	**A2**	**A3**	**A4**	**A5**	**A6**
C1	(0.02,0.023,0.027)	(1,1,1)	(0.52,0.52,0.83)	(0.5,1,1)	(0.025,0.031,0.031)	(0.066,0.076,0.083)
C2	(0.974,0.975,0.976)	(0.9997,0.9998,1)	(0.968,0.971,0.973)	(0.9990,0.9991,0.9994)	(0.995,0.996,0.9961)	(0.9953,0.9957,0.9963)
C3	(0.036,0.040,0.041)	(0.6,0.6,1)	(0.049,0.051,0.052)	(0.036,0.043,0.047)	(0.42,0.75,0.75)	(0.6,0.75,1)
C4	(0.0375,0.0377,0.054)	(0.75,0.75,1)	(0.0331,0.0337,0.034)	(0.034,0.035,0.037)	(0.5,0.545,0.6)	(0.46.0.545,0.6)
C5	(0.28,0.283,0.299)	(0.561,0.615,0.666)	(0.17,0.173,0.1772)	(0.53,0.54,0.55)	(0.914,0.969,1)	(0.301,0.301,0.32)
C6	0.977,0.978,0.978	(0.9990,0.9998,1)	(0.9753,0.9755,0.9755)	(0.9917,0.9924,0.9924)	(0.991,0.9925,0.9927)	(0.9785,0.9787,0.9797)

Table 6
Weighted normalized fuzzy decision matrix.

Criteria	Alternatives					
	A1	**A2**	**A3**	**A4**	**A5**	**A6**
C1	(0.020,0.023,0.081)	(1,1,3)	(0.52,0. 52, 2.49)	(0.5,1,3)	(0.025,0.031,0.093)	(0.066,0.076,0.249)
C2	(6.818,8.775,18.784)	(6.997,8.998,9)	(6.776,8.739,8.757)	(6.993,8.991,8.994)	(6.965,8.964,8.9649)	(6.967,8.961,8.966)
C3	(0.108,0.2,0.287)	(1.8,3,7)	(0.147,0.255,0.364)	(0.108,0.215,0.329)	(1.26,3.75,5.25)	(1.8,3.75,7)
C4	(0.1125,0.1885,0.378)	(2.25,3.75,7)	(0.099,0.1685,0.238)	(0.102,0.175,0.259)	(1.5,2.725,4.2)	(1.38,2.725,4.2)
C5	(1.96,2.547,2.691)	(3.927,5.535,5.994)	(1.19,1.557,1.593)	(3.71,4.86,4.95)	(6.398,8.721,9)	(2.107,2.709,2.88)
C6	(4.885,6.846,8.802)	(4.995,6.998,9)	(4.8765,6.828,8.779)	(4.958,6.946,8.931)	(4.955,6.9475,8.934)	(4.892,6.8509,8.8173)

Table 7
Distances between two alternatives for various attributes.

C1	X11	X21	X31	X41	X51	X61
X11	–	(3.23,0)	(0.173,0)	(3.115,0)	(0.0152,0)	(0.182,0)
X21		–	(0,3.06)	(0,0.5)	(0,3.215)	(0,3.04)
X31			–	(2.944,0)	(0,0.1597)	(0.02778,0)
X41				–	(0,3.10)	(0,2.934)
X51					–	(0.1674,0)
X61						–
C2	X11	X21	X31	X41	X51	X61
X11	–	(0.3583,0)	(0,0.061)	(0.3483,0)	(0.3,0)	(0.299,0)
X21		–	(0,0.4182)	(0,0.01)	(0,0.0584)	(0,0.0585)
X31			–	(0.408,0)	(0.3599,0)	(0.3597,0)
X41				–	(0,0.04857)	(0,0.04857)
X51					–	(0.003769,0)
X61						–
C3	X11	X21	X31	X41	X51	X61
X11	–	(7.476,0)	(0.10234,0)	(0.044,0)	(6.209,0)	(7.78,0)
X21		–	(0,7.369)	(0,7.424)	(0,1.979)	(0.75,0)
X31			–	(0.0659,0)	(6.109,0)	(7.68,0)
X41				–	(6.167,0)	(7.737,0)
X51					–	(1.831,0)
X61						–
C4	X11	X21	X31	X41	X51	X61
X11	–	(7.817,0)	(0,0.142)	(0,0.12)	(4.972,0)	(4.759,0)
X21		–	(0,7.948)	(0,7.927)	(0,3.0745)	(0,3.106)
X31			–	(0.022,0)	(4.9189,0)	(4.886,0)
X41				–	(4.898,0)	(4.865,0)
X51					–	(0,0.12)
X61						–
C5	X11	X21	X31	X41	X51	X61
X11	–	(4.869,0)	(0,1.669)	(3.6764,0)	(9.88,0)	(0.289,0)
X21		–	(0,6.533)	(0,1.262)	(5.029,0)	(0,4.582)
X31			–	(5.341,0)	(11.545,0)	(1.955,0)
X41				–	(6.208,0)	(0,3.388)
X51					–	(0,9.592)
X61						–
C6	X11	X21	X31	X41	X51	X61
X11	–	(0.2727,0)	(0,0.0304)	(0.179,0)	(0.1806,0)	(0.0175,0)
X21		–	(0,0.303)	(0,0.0939)	(0,0.092)	(0,0.256)
X31			–	(0.2089,0)	(0.2108,0)	(0.047,0)
X41				–	(0,0.0045)	(0,0.162)
X51					–	(0,0.164)
X61						–

Table 8
Concordance set of criteria.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	(0)	(2,4,5,6)	(4)	(0)	(0)
A2	(1,2,3,4,5,6)	–	(1,2,3,4,5,6)	(1,2,3,4,5,6)	(1,2,3,4,6)	(1,2,4,5,6)
A3	(1,3)	(0)	–	(0)	(1)	(0)
A4	(1,2,3,5,6)	(0)	(1,2,3,4,5,6)	–	(1,2,6)	(1,2,5,6)
A5	(1,2,3,4,5,6)	(5)	(2,3,4,5,6)	(3,4,5)	–	(4,5,6)
A6	(1,2,3,4,5,6)	(3)	(1,2,3,4,5,6)	(3,4)	(1,2,3)	–

formers. Taking these fuzzy ratings into account a decision matrix is prepared using Eq. (1). This decision matrix is shown in Table 4.

After preparation of fuzzy decision matrix, a normalized fuzzy decision matrix is prepared using Eqs. (2)–(4). This normalized fuzzy decision matrix is shown in Table 5.

Now, each element of the normalized fuzzy decision matrix is multiplied by the weight of corresponding criteria to find the weighted normalized fuzzy decision matrix. The weighted normalized fuzzy decision matrix is prepared using Eq. (5) and shown in Table 6.

The distances between two alternatives are calculated using the Eq. (5) for all the criteria to find the concordance and discordance sets. These distances are shown in Table 7.

Using the assumptions as mentioned in step 5 in research methodology section, a concordance set of criteria is prepared as shown in Table 8.

Now, the fuzzy weights of the importance of all the criteria in one set are added and the sum is shown in Table 9 as concordance matrix. For example, the element of A1–A3 cell can be calculated

Table 9
Concordance Matrix.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	(0)	(22,30,34)	(3,5,7)	(0)	(0)
A2	(26,36,44)	–	(26,36,44)	(26,36,44)	(19,27,35)	(23,31,37)
A3	(4,6,10)	(0)	–	(0)	(1,1,3)	(0)
A4	(23,31,37)	(0)	(26,36,44)	–	(13,17,21)	(20,26,30)
A5	(26,36,44)	(7,9,9)	(25,35,41)	(13,19,23)	–	(15,21,25)
A6	(26,36,44)	(3,5,7)	(26,36,44)	(6,10,14)	(11,15,19)	–

Table 10
Boolean matrix E.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	**0**	1	0	0	0
A2	1	–	1	1	1	1
A3	0	0	–	0	0	0
A4	1	0	1	–	0	1
A5	1	0	1	1	–	1
A6	1	0	1	0	0	–

as the summation of weights of concordance set of criteria C_2, C_4, C_5, and C_6, i.e.

$$((7, 9, 9) \oplus (3, 5, 7) \oplus (7, 9, 9) \oplus (5, 7, 9) = (22, 30, 34))$$

The Boolean matrix E is prepared from the concordance matrix. To prepare the Boolean matrix the average value of all the fuzzy numbers in concordance matrix is calculated as $\tilde{\bar{c}} = \left(\frac{390}{30}, \frac{540}{30}, \frac{660}{30}\right) = (13.18, 22)$ using Eq. (9). If the value $\tilde{\bar{c}}$is greater than the individual value in concordance matrix the value in Boolean matrix will be 1 otherwise 0 as shown in Table 10. For Example, the element of A1-A3 in Table 9 is (22,30,34) which is greater than the value of $\tilde{\bar{c}} = (13.18, 22)$; thus, the value of the element of the cell A1A3 in Boolean matrix E will be 1.

Using the assumptions in the step 7 and Eq. (11), a discordance set of criteria and discordance matrix are prepared as shown in Tables 11 and 12, respectively. For example, an element of A1-A3 discordance set of criteria is (C_1,C_3) in discordance matrix is calculated as

$$= \frac{\max(|0 - 0.173|,\ |0 - 0.1023|)}{\max(|0 - 0.173|,\ |0 - 0.061|,\ |0 - 0.1023|,\ |0 - 0.142|,\ |0 - 1.669|,\ |0 - 0.0304|)} = \frac{0.173}{1.669} = 0.1 \quad .$$

The discordance matrix is transformed into the Boolean matrix F as shown in Table 13. To transform the elements of the discordance matrix into Boolean form, the average value of all the elements of the discordance matrix is calculated as $\bar{d} = \frac{17.05}{30} = 0.568$ using Eq. (12). If the value of $\bar{d}$ is less than the individual value of the elements in discordance matrix, the value in Boolean matrix F will be 1 otherwise 0. For Example, the element of A1-A3 in Table 11 is 0.1 which is less than the value of $\bar{d}$(0.568); thus the value of the element of the cell A1-A3 of Boolean matrix F will be 1.

A global matrix G is prepared by multiplying the elements of Boolean matrix E with the corresponding elements of Boolean matrix F using equation (14). The global matrix G is shown in Table 14.

Table 12
Discordance matrix.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	1	0.1	1	1	1
A2	0	–	0	0	1	0.16
A3	1	1	–	1	1	1
A4	0.03	1	0	–	1	1
A5	0	0.64	0.01	0.49	–	0.19
A6	0	1	0	0.43	1	–

Table 13
Boolean matrix F.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	0	1	0	0	0
A2	1	–	1	1	0	1
A3	0	0	–	0	0	0
A4	1	0	1	–	0	0
A5	1	0	1	1	–	1
A6	1	0	1	1	0	–

Table 14
Global matrix G.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	0	1	0	0	0
A2	1	–	1	1	0	1
A3	0	0	–	0	0	0
A4	1	0	1	–	0	0
A5	1	0	1	1	–	1
A6	1	0	1	0	0	–

Table 15
Rank of Telecom Service Providers based on fuzzy ELECTRE.

Alternatives	Non-Dominant service providers	Rank
A1	A3	3
A2	A1,A3,A4,A6	1
A3	–	-
A4	A1,A3	2
A5	A1,A3,A4,A6	1
A6	A1,A3	2

On the basis of the global matrix G, the ranks are provided to the different alternatives, i.e. telecom service providers as shown in Table 15. The Boolean value 1 indicates the outranking, for example, in row 1, A1 outranks only A3; similarly, in row 2, A2 outranks A1, A3, A4, A6. A1 outranks A3 means A1 is performing better than the A3.

Table 11
Discordance set of criteria.

Alternatives	A1	A2	A3	A4	A5	A6
A1	–	(1,2,3,4,5,6)	(1,3)	(1,2,3,5,6)	(1,2,3,4,5,6)	(1,2,3,4,5,6)
A2	(0)	–	(0)	(0)	(5)	(3)
A3	(2,4,5,6)	(1,2,3,4,5,6)	–	(1,2,3,4,5,6)	(2,3,4,5,6)	(1,2,3,4,5,6)
A4	(4)	(1,2,3,4,5,6)	(0)	–	(3,4,5)	(3,4)
A5	(0)	(1,2,3,4,6)	(1)	(1,2,6)	–	(1,2,3)
A6	(0)	(1,2,4,5,6)	(0)	(1,2,5,6)	(4,5,6)	–

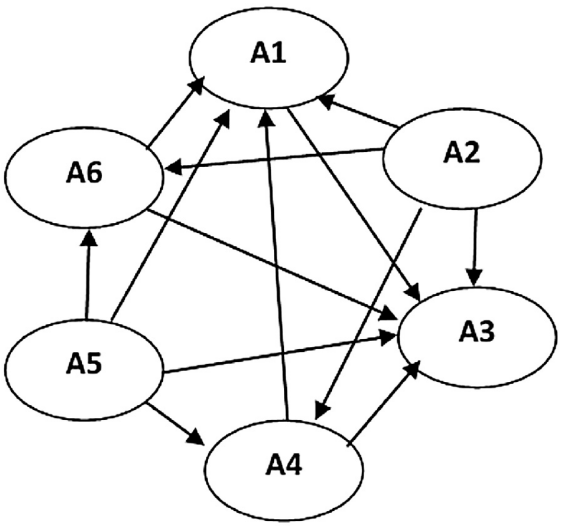


Fig. 3. ELECTRE outranking digram of the alternatives.

Using Table 15, an ELECTRE outranking diagram of alternatives is prepared as shown in Fig. 3.

Airtel and Reliance Communications are in the top rank in terms of the operational performance. Therefore, both of them are kept at the top rank. The accumulated downtime of Aircel and Reliance Communications are very high, but accumulated downtime among the performance parameters has least important. Call setup success rate of Airtel. Idea, Reliance Communications, and Vodafone are very high, whereas Aircel and MTNL have a lesser call set up success rate than that of the other service providers. This parameter is the most important parameter and Airtel and Idea have got the highest value in terms of call setup success rate. Aircel has highest channel congestion. Idea has second highest and MTNL has third highest channel congestion. Airtel, Reliance Communications, and Vodafone have least channel congestion. Airtel has least TCH congestion. MTNL has highest TCH congestion. Idea and Aircel have second and third highest TCH congestion, respectively. The call drop rate of MTNL is very high which is undesirable. Airtel and Vodafone have less value of the call drop rate. Aircel and Vodafone have second and third highest call drop rate, respectively. The call drop rate is a most influential factor to improve the operational performance. TRAI has taken a serious note of the call drop rate in India. Airtel, Idea, and Reliance Communications have the highest value of connecting with good voice level.

In terms of performance and subscribers, Airtel has the highest value and Reliance communications has the second rank, but using fuzzy ELECTRE both have the same ranking i.e. at the top level based on the operational performance. Similarly, Idea and Vodafone have the same ranking and they are on the second level. Aircel has the lowest performance level, whereas MTNL is outranked by the other service providers using fuzzy ELECTRE.

5. Conclusion

This paper has tried to illustrate a comprehensive approach for ranking and selection of cellular mobile telephone service providers in the Delhi working area. It is realized that the choice of the best service provider is a critical issue of the subscription of the services and it can be solved by the outranking approach. The fuzzy extended ELECTRE approach incorporates the linguistic vagueness in decision making. It also outranks the service providers which are a poor performer. A number of parameters and sub-parameters have been taken into consideration before making the decision about the selection of the location. We ought to recognize every last trace of the vital criteria such as network availability (BTS accumulated downtime), connection accessibility (call setup success rate, channel congestion, and TCH congestion), and connection retainability (call drop rate and connection with good voice quality). The service providers should have lower values of BTS accumulated downtime, channel, and TCH congestions and call drop rate; whereas they should have higher values of call set up success rate and connection with good voice quality.

The Delhi working area has the highest teledensity in India. Thus, the proposed approach can also be considered as the benchmarking model for operational performance of cellular mobile telephone service providers. The same model can also be used in another area for the similar type of the analysis. But, the analysis result can vary from one decision maker to another because the importance rating of the criteria is based on the perception of the decision makers. Nowadays, the call drop rate is a big issue in India. TRAI has considered it as a serious issue. It has been observed that MTNL has highest call drop rate and lower call setup success rate as per TRAI report. This is the reason that MTNL is outranked by all the other telecom service providers in Delhi working area. MTNL should resolve these issues to improve its performance since the telecom industry is growing at a fast rate and are highly competitive in India.

In future, this work will be extended to the larger area, i.e. at the country level. Some more operational parameters can be also incorporated for comprehensive analysis in future. After observing the lagging operational parameters, a company may think for further improvement in terms of respective parameters.

Acknowledgment

Authors acknowledge the valuable suggestions of Editor and Reviewers of the journal to improve the quality and content of this research work.

References

[1] Shiu-Wan Hung, Wen-Min Lu, A comparative study of the performance measurement in global telecom operators, Total Qual. Manage. 18 (2007) 1117–1132.
[2] V. Nigam, T. Thakur, V.K. Sethi, R.P. Singh, Benchmarking of Indian mobile telecom operators using DEA with sensitivity analysis, Benchmark.: Int. J. 19 (2012) 219–238.
[3] N.H. Leff, Social benefit-cost analysis and telecommunications investment in developing countries, Inf. Econ. Policy 1 (1984) 217–227.
[4] R.M. Debnath, R. Shankar, Benchmarking telecommunication service in India, Benchmark.: Int. J. 15 (2008) 584–598.
[5] S. Kumar, Y.S. Kumar, Evaluation of comparative performance of telecom service providers in India using TOPSIS and AHP, Int. J. Bus. Excellence. 6 (2013) 192–213.
[6] TRAI (April-June, 2015). Audit & Assessment of Quality of Service Of Cellular Mobile Telephone Service For Telecom Regulatory Authority of India North Zone –Delhi Service Area (April 2015–June 2015). Available at http://www.trai.gov.in/WriteReadData/WhatsNew/Documents/PR-37-Apr-15. pdf Accessed on November 5, 2015.
[7] TRAI, Annual Report 2013–14, Telecom Regulatory Authority of India, 2014, Available at: http://www.trai.gov.in/WriteReadData/UserFiles/Documents/AnuualReports/TRAI-Annual-Report-(English)=01042015.pdf, Accessed on 5th November 2015.
[8] A. Kumar, R. Shankar, R.M. Debnath, Analyzing customer preference and measuring relative efficiency in telecom sector: a hybrid fuzzy AHP/DEA study, Telematics Inf. 32 (2015) 447–462.
[9] V. Saxena, T. Thakur, R.P. Singh, Evaluating the performance of mobile telecom operators in India, Int. J. Simul. Syst. Sci. Technol. 10 (2009) 58–63.
[10] H.-C. Tsai, C.-M. Chen, G.W. Tzeng, The comparative productivity efficiency for global telecoms, Int. J. Prod. Econ. 103 (2006) 509–526.
[11] D.I. Giokas, G.C. Pentzaropoulos, Evaluating productive efficiency in telecommunications: evidence from Greece, Telecommun. Policy 24 (2000) 781–794.
[12] J. Zhu, Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets, Kluwer Academic Publishers, Boston, 2003.

[13] H.-H. Yang, C.-Y. Chang, Using DEA window analysis to measure efficiencies of Taiwan's integrated telecommunication firms, Telecommun. Policy 33 (2009) 98–108.
[14] M. Resende, Efficiency measurement and regulation in US telecommunications: a robustness analysis, Int. J. Prod. Econ. 114 (2008) 205–218.
[15] S. Joshi, Customer experience management: an exploratory study on the parameters affecting customer experience for cellular mobile services of a telecom company, Procedia – Soc. Behav. Sci. 133 (2014) 392–399.
[16] S.K. Majumdar, Does competition enhance performance? Entry policy and efficiency patterns in telecommunications markets, Info 12 (2010) 25–38.
[17] N. Yadav, Total interpretive structural modelling (TISM) of strategic performance management for Indian telecom service providers, Int. J. Prod. Perform. Manage. 63 (2014) 421–445.
[18] M. Bottero, V. Ferretti, J.R. Figueira, S. Greco, B. Roy, Dealing with a multiple criteria environmental problem with interaction effects between criteria through an extension of the Electre III method, Eur. J. Oper. Res. 245 (2015) 837–850.
[19] J. Żak, M. Kruszyński, Application of AHP and ELECTRE III/IV Methods to Multiple Level, Multiple Criteria Evaluation of Urban Transportation Projects. 18th Euro Working Group on Transportation, EWGT 2015, 14–16 July 2015, Delft, The Netherlands, Transp. Res. Procedia 10 (2015) (2015) 820–830.
[20] B. Ka, Application of fuzzy AHP and ELECTRE to China dry port location selection, Asian J. Shipping Logist. 27 (2011) 331–353.
[21] G.A. Montazer, H.Q. Saremi, M. Ramezani, Design a new mixed expert decision aiding system using fuzzy ELECTRE III method for vendor selection, Expert Syst. Appl. 36 (2009) 10837–10847.
[22] A.C. Tolga, A real options approach for software development projects using fuzzy ELECTRE, J. Multiple-Valued Logic Soft Comput. 18 (2012) 541–560.
[23] F. Zagonari, Choosing among weight-estimation methods for multi-criterion analysis: a case study for the design of multi-purpose offshore platforms, Appl. Soft Comput. 39 (2016) 1–10.
[24] A. Hatami-Marbini, M. Tavana, M. Moradi, F. Kangi, A fuzzy group ELETRE method for safety and health assessment in hazardous waste recycling facilities, Saf. Sci. 51 (2013) 414–426.
[25] R.W. Hamming, Error detecting and error correcting codes, Bell Syst. Tech. J. 26 (1950) 147–160.
[26] C.-T. Chen, Extensions of the TOPSIS for group decision-making under fuzzy environment, Fuzzy Sets Syst. 114 (2000) 1–9.
[27] D.-F. Li, J.-B. Yang, Fuzzy linear programming technique for multiattribute group decision making in fuzzy environments, Inf. Sci. 158 (2004) 263–275.
[28] E. Herrera-Viedma, O. Cordon, M. Luque, A.G. Lopez, A.M. Munoz, A model of fuzzy linguistic IRS based on multi-granular linguistic information, Int. J. Approx. Reason. 34 (2–3) (2003) 221–239.
[29] Chia-Chi Sun, A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods, Expert Syst. Appl. 37 (12) (2010) 7745–7754.
[30] W. Tien-Chin, Tsung-Han Chang, Application of TOPSIS in evaluating initial training aircraft under a fuzzy environment, Expert Syst. Appl. 33 (4) (2007) 870–880.
[31] A. Awasthi, S.S. Chauhan, S.K. Goyal, A fuzzy multicriteria approach for evaluating environmental performance of suppliers, Int. J. Prod. Econ. 126 (2) (2010) 370–378.
[32] M. Dursun, E. Ertugrul Karsak, A fuzzy MCDM approach for personnel selection, Expert Syst. Appl. 37 (6) (2010) 4324–4330.
[33] Chih-Hsuan Wang, Hui-Shan Wu, A novel framework to evaluate programmable logic controllers: a fuzzy MCDM perspective, J. Intell. Manuf. 27 (2) (2016) 315–324.
[34] P. Kumar, R.K. Singh, P. Sinha, Optimal site selection for a hospital using a fuzzy extended ELECTRE approach, J. Manage. Anal. 3 (2) (2016) 115–135.
[35] P. Kumar, R.K. Singh, R.K. Vaish, Suppliers' green performance evaluation using fuzzy extended ELECTRE approach, Clean Technol. Environ. Policy (2016) 1–13, http://dx.doi.org/10.1007/s10098-016-1268-y.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/317106456

A Simple and Efficient Method for Removal of Phenolic Contaminants in Wastewater Using Covalent Immobilized Horseradish...

Article · February 2017

DOI: 10.17265/2161-6221/2017.1-2.004

CITATIONS
0

READS
11

6 authors, including:



Rajesh Ahirwar
National Institute for Research in Environmental Health
21 PUBLICATIONS **35** CITATIONS

SEE PROFILE

All content following this page was uploaded by Rajesh Ahirwar on 25 May 2017.

The user has requested enhancement of the downloaded file. All in-text references underlined in blue are added to the original document and are linked to publications on ResearchGate, letting you access and read them immediately.

Journal of Materials Science and Engineering B 7 (1-2) (2017) 27-38
doi: 10.17265/2161-6221/2017.1-2.004

A Simple and Efficient Method for Removal of Phenolic Contaminants in Wastewater Using Covalent Immobilized Horseradish Peroxidase

Rajesh Ahirwar[1,2,6], Jai G. Sharma[3], Bhanumati Singh[4], Krishan Kumar[5], Pradip Nahar[1,2] and Saroj Kumar[1,3]

1. CSIR-Institute of Genomics and Integrative Biology, Delhi 110007, India

2. Academy of Scientific and Innovative Research, CSIR-Institute of Genomics and Integrative Biology, Delhi 110007, India

3. Department of Biotechnology, Delhi Technological University, Delhi 110042, India

4. Department of Biotechnology, Bundelkhand University, Jhansi, UP 284128, India

5. Department of Chemistry, Motilal Nehru College, University of Delhi, Delhi 110021, India

6. Present address: Department of Biochemistry, ICMR-National Institute for Research in Environmental Health, Kamla Nehru Hospital Building, Bhopal 462001, India

Abstract: In the light of the increasing burden of pollutants in major rivers and stringent environmental legislation, adaptation to eco-friendly treatment approaches is desperately required to decontaminate industrial effluents before its discharge to rivers and other fresh water-bodies. In present study, we have designed a simple and efficient method for removal of phenols from effluent wastewater using an immobilized preparation of HRP (horseradish peroxidase). The enzyme was isolated in bulk amount from the roots of the *Armoracia rusticana* and covalently immobilized to polycarbonate supports using a photolabile linker FNAB (1-fluoro-2-notro-4-azidobenzene). The immobilized enzyme showed enhanced storage stability and reusability. The immobilized HRP was subsequently used for degradation of phenols in sewage and spiked wastewater. The phenol content of spiked wastewater was reduced to 93% in the 3 L reactor following treatment with immobilized HRP and H_2O_2. The improvement in the quality of water upon treatment was reflected by the changes in pH, conductivity, TDS (total dissolved salts) and biodegradation of organic contents as indicated by 77% and 87% reduction in COD (chemical oxygen demand) and BOD (biochemical oxygen demand) respectively in the analyzed sample.

Key words: Phenolic wastewater, enzymatic treatment, horseradish peroxidase, covalent immobilization.

1. Introduction

Phenolic compounds are common contaminants of water bodies connected with the industrial effluents from petrochemical, organic chemical, resins, plastic and textile industries [1-4]. Majority of phenols are toxic substances, and some are known or suspected carcinogens classified as hazardous wastes [5]. As per ISI (Indian standard institution) limits, the permissible limit of phenol/phenolic compounds in inland water resources should not cross 0.001 mg/L (permissible limit in the absence of alternate source is 0.002 mg/L). Phenol at a concentration level of as low as 0.005 mg/L imparts a characteristic smell upon chlorination, making it unsafe for drinking purposes. The effects of phenolic compounds on humans include the irritation of nose, throat and eyes, digestive difficulties, nervous problems, headaches and skin burns. Prolonged ingestion of phenol at concentration levels ranging 10-240 mg/L can cause mouth sore, diarrhea and impaired vision [5]. The industrial effluents, which may contaminate the rivers and other water resources, carry phenolic content in the range of 200 to 2,000 mg/L. If released untreated to rivers, this large

Corresponding author: Saroj Kumar, Dr., young scientist, research fields: biotechnology, bionanotechnology, bioconjugation chemistry. E-mail: sarojkumar@dtu.ac.in.

concentration of phenols in wastewater can cause enormous damage to both the aquatic and human lives. Therefore, removal of phenol and its derivatives from contaminated water before discharge into river or other water bodies is strongly recommended to save aquatic organisms and preserve the environmental quality [6, 7]. Various conventional methods have been applied to remove phenols from wastewater including steam distillation, liquid-liquid extraction, adsorption, solid-phase extraction, wet air oxidation, catalytic wet air oxidation and biodegradation [8]. With further advancement in technologies, newer methods like electrochemical oxidation, photo-oxidation, coronation, UV/H_2O_2, Fenton reaction, membrane processes and enzymatic treatment have been used for phenol degradation [2-4, 9, 10]. Still, many of these methods have disadvantages like low efficiency, high cost, partial removal and formation of hazardous by-products [11-13]. Hence, the need to develop an efficient, simple and scalable method for removal of phenolic contaminants from wastewater still persist. Enzymes such as peroxidases obtained from horseradish, soybean and other plant sources have been used to remove phenols from waste effluents in the presence of hydrogen peroxide [14-17]. In presence of H_2O_2, the native HRP (horseradish peroxidase) is oxidized to an active intermediate (HRPi) that catalyzes oxidation of aromatic compounds, producing a free radical (AH•) and returns to its native state (HRP). The released free radicals, in bulk solution, form polyaromatic products, which are water-insoluble and remove as precipitate [9, 17, 19]. As shown in the following reaction steps Eqs. (1)-(3).

$$\text{HRP} + H_2O_2 \rightarrow \text{HRPi} + H_2O \quad (1)$$

$$\text{HRPi} + \text{Aromatic compound } (AH_2) \rightarrow \text{HRPii} + \text{Free Radical } (AH\bullet) \quad (2)$$

$$\text{HRPii} + \text{Aromatic compound } (AH_2) \rightarrow \text{HRP} + \text{Free Radical } (AH\bullet) + H_2O \quad (3)$$

The major issue with peroxidases is the reduction in phenol removal efficiency at low concentrations of the enzyme due to the inactivation of the enzyme. Use of higher concentrations of purified enzyme can lead to tremendous increase in capital investment for bulk purification facility. Also, the separation of enzyme from treated water is not possible without its immobilization on some solid supports [20]. As the immobilization of enzymes may solve such limitations, several approaches for covalent and non-covalent immobilization of peroxidase enzyme to solid matrices are reported in the literatures [21-25]. The covalent approach is usually preferred as that it restricts the loss of enzyme, allowing more enzymes per unit amount of degradable phenolic contaminants.

To develop a simple and scalable method for enzymatic removal of phenolic contaminants from wastewater effluents, we have immobilized HRP on PC (polycarbonate) discs and used it to degrade the phenolic contaminants in sewage water. Further, the effect of PEG (polyethylene glycol) addition on HRP activity during dephenolization was studied. Finally, the immobilized preparation of HRP was used in a lab-made wastewater treatment cell (bioreactor; 3 L capacity) to study its effectiveness in lowering the phenolic contaminants in the collected sewage wastewater samples. A comparison of treated and untreated wastewater was made in terms of change in physical parameters like pH, salinity, conductivity, TDS (total dissolved salts) and chemical parameters including residual phenolic contaminants, ammonia (NH_3), sodium (Na^+) and chloride (Cl^-) ions, nitrates (NO_3^-) and nitrites (NO_2^-), BOD (biochemical oxygen demand) and COD (chemical oxygen demand).

2. Experimental Setup

2.1 Chemicals and Reagents

Phenol, *p*-chlorophenol, potassium ferricyanide, 4-aminoantipyrine (4-AAP), ammonium sulphate, polyethylene glycol (PEG; average molecular weight 3,350) were purchased from SRL (scientific research laboratory), India Pvt Ltd. Hydrogen peroxide (H_2O_2; 30%, v/v), was purchased from Merck Life Science

Pvt. Ltd. Bengaluru, India. Photolinker FNAB was synthesized using 4-fluoro-3-nitroaniline (Sigma-Aldrich, Germany) as reported earlier in Ref. [26]. Photo irradiation of polycarbonate discs was carried out at a wavelength of 365 nm in a UV Stratalinker, model-2400, (Stratagene, USA). Polycarbonate sheets were purchased from the local market (New Delhi, India). All the experiments were carried out in triplicates. The absorbance of samples was recorded in an ELISA reader (BioRad laboratories India).

2.2 Isolation and Purification of HRP

Horseradish peroxidase was isolated and purified from the roots of horseradish (*Armoracia rusticana*) according to previously reported procedure [27]. Briefly, the roots of horseradish were crushed to collect the homogenate. Crude HRP was obtained from the collected homogenate using ammonium sulphate precipitation (90% saturation) and dialysed in 5 mM acetate buffer (pH 4.4) for overnight. Next, the collected crude enzyme was fractioned on an ion exchange column (Whatman's carboxymethylcellulose CM52) to obtain partially purified enzyme. Further purification was carried out using size-exclusion chromatography on HiPrep Sephacryl S-100 HR column (GE Healthcare Life Science, India).

2.3 Measurement of HRP Activity

Measurement of HRP activity is based on the mechanism of HRP action. HRP combines with H_2O_2 (hydrogen peroxide) to form an intermediate HRP-H_2O_2 complex, which can oxidize a wide variety of hydrogen donors.

$$\text{Donor} + H_2O_2 \rightarrow \text{Oxidized Donor} + 2\ H_2O \quad (4)$$

We measured the enzyme activity of the purified HRP using modified Worthington method [28]. During the HRP biocatalysis, the hydrogen donor couple phenol/4-aminoantipyrine was converted to a pink-coloured reagent with maximum absorption (λ_{max}) at 510 nm, which was used in the assay of HRP activity. Briefly, the method involves incubation of 0.1 mL of HRP (in 0.1 M phosphate buffer, pH 7.0) with 1.5 mL of 0.17 mM H_2O_2 (in 0.1 M phosphate buffer, pH 7.0) and 1.4 mL of phenol/ 4-aminoantipyrine solution (0.17 M phenol with 2.5 mM 4-AAP in MilliQ water) for 1-5 minutes of time period. Increase in absorbance (λ = 510 nm) was recorded at different time points during the 1-5 min incubation period. One unit of HRP activity (U) is defined as the amount of HRP required to hydrolyze 1 µmol of H_2O_2 in 1 min at 25 ºC and pH 7.0. Specific activity can be calculated as unit activity per mg HRP.

Unit activity was measured as:

$$U/mL = \frac{(\Delta A510/\min test - \Delta A510/\min control) \times 3 \times df}{\varepsilon \times 0.1} \quad (5)$$

where 3 is the total volume of assay solution, 0.1 mL is the volume of enzyme and df is the dilution factor. The molar extinction coefficient (ε) of chromogenic dye (phenol/4-aminoantipyrine solution) is 7,100/M cm [29].

2.4 Activation of Polycarbonate Disc

Circular polycarbonate discs (total surface area = 50-180 cm^2) were prepared for covalent immobilization of HRP by covering them with the ethanolic solution of FNAB (1.82 mg FNAB per 30 µL ethanol) followed by the evaporation of ethanol in a dark fume hood to form a layer of FNAB over the PC discs. The discs were then irradiated with UV light at 365 nm in a UV-Stratalinker for 15 min [30]. After this, the unreacted FNAB on PC-discs was removed by ethanol wash and the discs were air-dried to get the FNAB activated disc ready for covalent immobilization.

2.5 Immobilization of HRP onto Activated Polycarbonate Discs

The activated PC discs were taken in petri dishes and loaded with different amounts (0.25-4 U/mL) of HRP in phosphate-buffered saline (0.1 M PBS, pH 7.0). Immobilization was carried out in a laboratory

incubator by 45 min incubation at 50 °C. After incubation, the discs were washed with washing buffer (0.1 M PBS, pH 7.0 with 0.1% Tween 20) and stored at 4 ºC. The unbound HRP from activated PC discs was collected to determine the immobilization capacity and the concentration of residual HRP in solution was determined by the method of bicinchoninic acid [31]. The amount of the immobilized enzyme was calculated as:

$$Q = \frac{(C_0 - C)V}{A} \quad (6)$$

where, Q is the amount of HRP immobilized per unit area of polycarbonate discs (mg/cm^2), C_0 and C are the concentrations (mg/mL) of HRP before and after the immobilization. V is the volume of the used HRP solution, and A is the surface area of the polycarbonate discs.

2.6 Effect of pH and Temperature on the Activity of Immobilized HRP

The activity of immobilized and free HRP was checked at different pH (3 to 6 in 0.1 M citrate-phosphate buffer and 7 to 9 in 0.1 M phosphate buffer). The effect of temperature on the enzymatic activity of immobilized and free HRP was determined by measuring the residual activity after 60 min incubation in 0.1 M phosphate buffer, pH 7.0 at various temperatures (27 °C-60 °C).

2.7 Degradation of the Phenolic Compounds in Spiked MilliQ Using PC-HRP

Oxidative degradation of phenol and p-chlorophenol using the immobilized HRP was studied in petri dishes containing 2 mL of 128 mg/L phenolic mixture (in MilliQ water) and PC-HRP disc (total surface area = 48 cm^2) in 7.75 mL of 0.1 M PBS, pH 7.0. The enzymatic reaction was initiated by adding 0.25 mL of 10 mM H_2O_2 after which the assembly was incubated for 3 h at room temperature. Degradation of phenolic compounds was monitored by withdrawing 200 µL aliquots from the reaction mixture and estimating the residual phenols by the colorimetric method as reported earlier [32]. Briefly, the assay was carried out by mixing 200 µL of collected phenolic sample with 25 µL of 83.4 mM $K_3Fe(CN)_6$ (in $NaHCO_3$) and 25 µL of 20.8 mM of 4-aminoantipyrine (in 0.25 M $NaHCO_3$). The reaction was incubated for 2-5 min, after which the absorbance was measured at 510 nm. The absorbance units provide an estimate of the non-degraded phenols.

Also, the protective effect of PEG on the enzymatic activity of immobilized HRP during dephenolization was studied. For this, the PC-HRP discs (total surface area = 48 cm^2) were loaded with 10 mL solution of the prepared phenol contaminated water (64 mg/L of each phenol and p-chlorophenol in MilliQ water) and incubated in presence of different amount of PEG (0.25, 0.5, 1.0, and 2.0 mg/mL) for 3 h at room temperature. After stipulated time, the residual phenol in each sample was measured and plotted as relative activity (HRP) against PEG concentrations.

2.8 Degradation of the Phenolic Compounds in Sewage Wastewater Using PC-HRP

An assembly for enzymatic degradation of phenolic contaminants in sewage water was prepared by attaching the HRP-immobilized PC-discs to a rod (shaft) fitted to an inverted mixer. The rod along with its inverted-mixer was mounted in a glass beaker of 3 L capacity. The beaker was filled with sewage water collected from the local drain (Najafgarh Drain; collection point- Kingsway Camp, Delhi, India). The enzymatic reaction was initiated by adding H_2O_2 to a final concentration of 0.25 mM. This followed the addition of optimised concentration of PEG. The sample was kept stirring (100-120 RPM) with the help of fitted PC-HRP discs. Total phenol in sewage wastewater before and after treatment (at 1 h, 2 h, 3 h and 14 h intervals) was measured using the colorimetric method [32]. Various physiological and chemical parameters of the collected sewage water

like the pH, total dissolved salts, conductivity, salinity, phenolic contaminants, ammonia (NH_3), sodium ion (Na^+), chloride ion (Cl^-), nitrates (NO_3^-) and nitrites (NO_2^-), BOD and COD were also measured before and after treatment. Total dissolved salts, pH, salinity and conductivity measurements were performed on Seven Excellence™ pH/conductivity/ion meter (model- S400, METTLER TOLEDO, Switzerland). The nitrates and nitrites were detected using the spectrophotometric method as reported elsewhere shown in Ref. [33]. Determination of ammonia was performed using the Orion AQUAfast AC4010 ammonia ultra low range auto-test ampoules with the Orion AQUAfast AQ4000 colorimeter (Thermo Scientific). Analysis of sodium and chlorine was made using kits availed from HACH, USA. BOD and COD measurements were performed as per "Standard Analytical Procedures for Water Analysis" set up by Central Water Commission, Govt. of India [34].

2.9 Reusability and Storage Stability of Immobilized HRP

The reusability of PC-immobilized HRP for phenolic degradation was investigated in a 3 L bioreactor. During each reusability experiment, the bioreactor system was operated for 2 h. The storage stability of immobilized HRP was determined after 0, 2, 4, 8 and 16 days storage in a sealed polyethylene bag in a refrigerator at 4 °C. The activity of immobilized HRP was expressed as percentage of its residual activity with respect to their initial activity.

3. Results and Discussion

3.1 Preparation of HRP Immobilized PC-discs

We have isolated and purified HRP enzyme from the roots of *Armoracia rusticana* using a low-cost setup. A yield of 80 mg purified HRP was obtained from 1 Kg horseradish (Fig. 1). The calculated specific activity of the purified lyophilized HRP obtained after ion-exchange and size exclusion chromatography was 100 U/mg. The polycarbonate discs for covalent immobilization of HRP were prepared by activating their surfaces using an aryl azide-based heterobifunctional crosslinker, FNAB. The linker coated discs were exposed to UV light to generate highly reactive nitrene, which attaches covalently to PC surface through C-H bond insertion [26, 35]. This activated polycarbonate having an active fluoro group enables bio-conjugation of HRP through the amino group of the enzyme when incubated at an elevated temperature of 40-50 ºC [36].

Further, we determined the immobilization capacity of HRP over PC discs by measuring the residual protein in left over unreacted HRP collected from immobilization experiments in which varying concentrations (0.25, 0.5, 1, 2, 4 U/mL) of HRP (specific activity = 100 U/mg) were immobilized onto the activated PC discs (total surface area = 1 cm^2). The immobilization capacities were calculated using Eq. (6). An optimal immobilization of 1.0 U/cm^2 was achieved. This suggests a PC disc of radius 1 meter would require 3.14 mg purified HRP (sp. Activity = 100 U/mg), which is considerably a cost-effective amount of enzyme.

Furthermore, the enzymatic degradation of o-phenylenediamine by the PC-HRP discs showed that HRP retain its enzymatic activity after immobilization

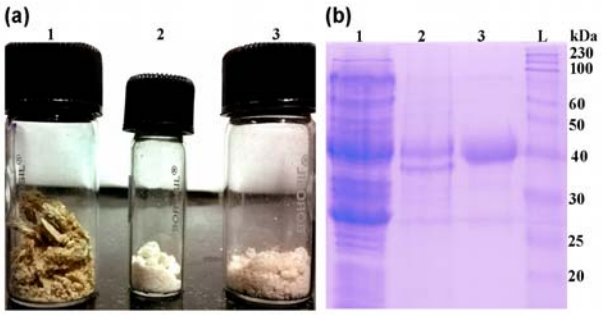


Fig. 1 Isolation and purification of HRP from the roots of horseradish. (a) Crude HRP a1 is obtained from horseradish homogenate after ammonium precipitation and purified on ion-exchange a2 and size-exclusion a3 columns. (b) SDS PAGE of HRP in its crude form (lane 1), after ion-exchange chromatography (lane 2) and after size-exclusion chromatography (lane 3). L represents the pre-stained protein ladder of size 10-230 kDa (NEB).

(Fig. 2). Control assay performed with untreated PC discs carrying physically adsorbed HRP produced an insignificant absorbance reads, suggesting an absolute requirement of chemical linkers for enzyme immobilization onto PC discs. Polycarbonate was taken as model support for enzyme immobilization because of its stability (both thermal and mechanical) and chemical inertness.

3.2 Optimizing the Reaction Parameters for Maximum Activity of the Immobilized HRP

The effect of pH on the activity of free and immobilized HRP was studied at pH 3-9. Citrate-phosphate buffer, 0.1 M was used to study the effect of pH in the range of 3-6, while phosphate buffer, 0.1 M was used for pH of 7-9. As shown in Fig. 3a, the maximum relative activity of immobilized and free enzyme was obtained at pH 7.0. Nevertheless, at any pH (except pH 7.0), the relative activity of immobilized HRP has remained higher than the free enzyme. This broadening in the pH profile of HRP upon immobilization could be explained by change in charge interactions upon formation of strong covalent bond between surface and enzyme [37, 38].

Similarly, the thermal stability of the immobilized and free HRP was determined by incubating them in PBS (0.1 M PBS, pH 7.0) for 60 min at 27 ºC, 40 ºC, 45 ºC, 50 ºC, 55 ºC and 60 °C, respectively. The residual activity of immobilized and free HRP was calculated with respect to its initial activity at room temperature of 27 ºC. As shown in Fig. 3b, both the immobilized and free enzyme retained similar activity at low temperatures of up to 50 °C, beyond which the activity of both immobilized and free HRP got decreased with increasing temperature. As expected, it was found that the PC-immobilized HRP retained approx 89% of its initial activity after 60 min incubation at 60 °C, whereas free enzyme retained only 73% of its initial activity at corresponding time and temperature. This suggests that the enzymatic activity and stability of HRP improve upon immobilization onto PC discs. Increase in thermal stability of HRP upon immobilization could be contributed to the same strong interactions between enzyme and support that affect the intra-molecular forces responsible for maintaining the conformation,

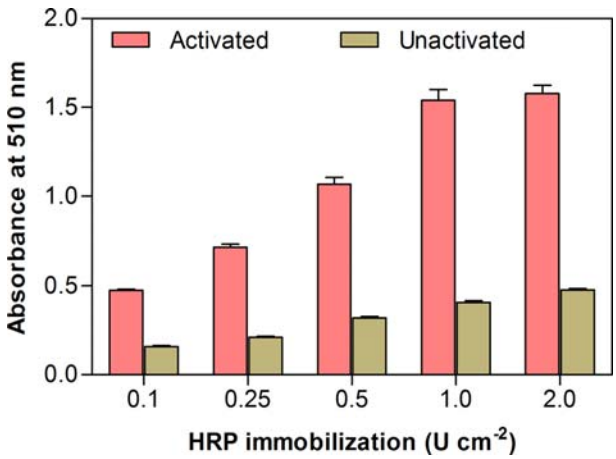


Fig. 2 Analysing the activity of immobilized HRP. Enzymatic conversion of colourless o-phenylenediamine (dihydrochloride) solution to an orange-brown coloured product using physically adsorb (un-activated) and chemically (using FNAB; activated) immobilized HRP on PC discs.

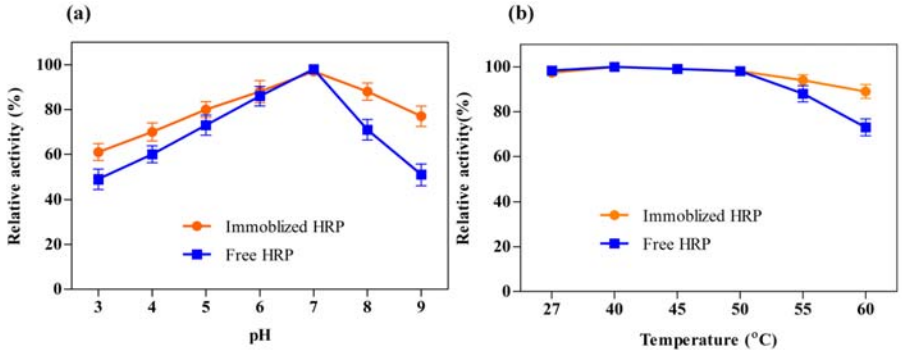


Fig. 3 (a) Effect of pH on the relative activity (against standard pH 7.4) of the immobilized and free HRP. (b) Effect of temperature on the relative activity (against standard temperature 27 ºC ± 2 ºC) of immobilized and free HRP.

thus the activity of the enzyme [39]. The observation of increased thermal stability and broader pH range upon immobilization imprints the idea of using PC-HRP in industrial setup.

3.3 Efficient Removal of Phenol and p-chlorophenol from Spiked MilliQ Water

To test the capacity of the developed PC-HRP system to remove phenols from water samples, we added 64 mg/L phenol and 64 mg/L p-chlorophenol to MilliQ water. The enzymatic degradation of phenols was studied at different time points; the PC-HRP disc with total surface area of 48 cm^2 (1.0 U/cm^2) degraded 80% phenols in first 30 min, 94% in 60 min and 96% phenols upon 3 h incubation (Fig. 4a). Like previous observations of 90% phenol degradation using purified HRP, we achieved an efficiency of 96% polymerization using the immobilized HRP [19, 40, 41].

The efficiency of phenol reduction can be enhanced by either increasing the enzyme concentration or by increasing the incubation time. However, longer incubations can result in enzyme inactivation, leading to low conversion [42]. Previous reports suggest that the presence of additives such as PEG enhance enzyme performance for the treatment of wastewater containing a range of phenol concentrations between 47-1,500 mg/L [43]. Accordingly, we used additive PEG (0-2 mg/mL) for its protective action on HRP during removal of phenol and p-chlorophenol. As shown in Fig. 4b, the HRP retains 93% of its initial activity in the presence of PEG at an optimal concentration of 1 mg/mL after 3 h incubation in the spiked phenolic water (128 mg/L phenols). Whereas, only 35% HRP activity was monitored in samples carrying no PEG. In HRP-catalysed phenolic reactions, the HRP inactivation mainly occurs as a result of sorption/occlusion by polymeric products (HRP adsorption on precipitated products occludes its active sites) [44] or upon heme-destruction (generated free radicals react and inactivate the heme-centre in HRP) [45]. This improvement in the efficiency of HRP by PEG addition during dephenolization is due to the formation of a protective layer in the vicinity of the active centre of HRP which restricts the attack of free phenoxy radicals and help in effectively suppressing the heme-destruction by co-dissolved PEG [44-46].

3.4 Removal of Phenolic Contaminants from Sewage Wastewater

The removal of phenolic contaminants from the collected sewage wastewater was performed in a 3 L glass beaker assembly as shown in Fig. 5. Due to lack of open sites of discharge of industrial effluents in Delhi, we have used water samples collected from local drain, which majorly contains the household sewage. Before and after treatment analysis of collected wastewater showed 91% reduction in phenolic

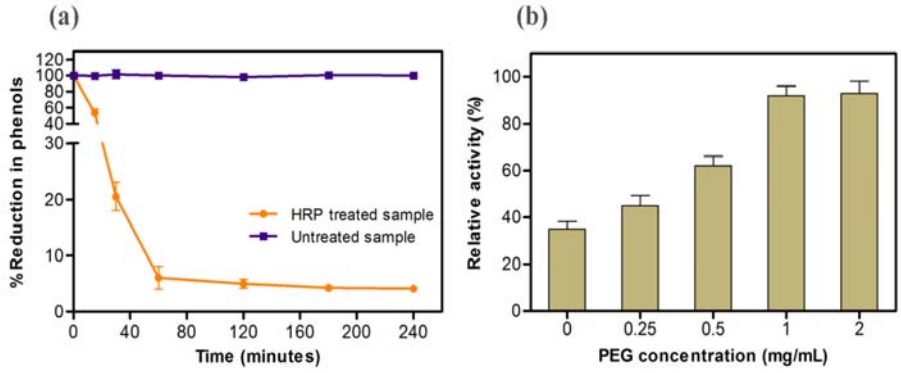


Fig. 4 Optimization of parameters for removal of phenolic contaminants from spiked water. (a) Time-dependent removal of phenol and p-chlorophenol from spiked water carrying total 128 mg/L of phenols using PC-HRP disc. (b) Optimization of PEG concentrations for removal of p-chlorophenol from contaminated water.

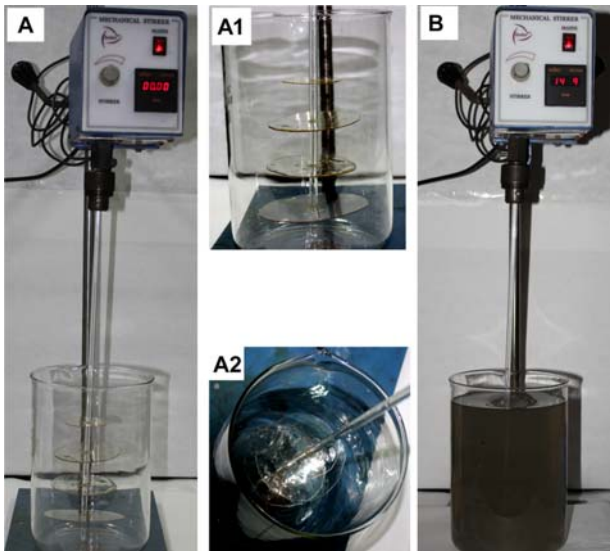


Fig. 5 Setup to remove phenols from contaminated wastewater. (A) The prepared empty assembly showing the PC-HRP disc attached to turbine of inverted rotor. A1 and A2 provide a close-up image of the turbine created using PC-HRP discs. (B) Phenol removal assembly with filled sewage water. Phenol degradation is initiated by adding 0.25 mM H_2O_2 to the filled effluent tank. The PC-HRP discs are rotated at a constant speed of 100-120 rpm to prevent settlement of added PEG (1 g/L). Samples are collected at 1, 2, 3, and 14 h intervals to monitor phenol degradation.

content. The initial concentration of phenols in the collected wastewater sample was calculated to 2-3 mg/L. As the collected wastewater majorly contains the household sewage, such a low amount of phenolic contamination was expected. Interestingly, after 3 h of oxidative degradation by PC-HRP, the phenolic contamination has reduced to 0.22 mg/L. Because the collected wastewater had low level of phenolic contaminants, to obtain an ideal sample for analysis, we have added phenol (10 mg/L), p-chlorophenol (10 mg/L), 2, 4-dichlorophenol (5 mg/L), catechol and 2, 4-dimethylphenol (5 mg/L) to a total concentration of 30 mg/L phenols in the collected wastewater sample. Enzymatic treatment of this spiked wastewater using the prepared assembly of PC-HRP resulted in phenol reduction by 81% in 1 h treatment, which further reduced to 93% in 2 h and 95% upon 3 h of treatment (Table 1). A further increase in the incubation time (e.g. overnight) showed no further reduction in phenolic contaminants. The improvement in the quality of water upon treatment

Table 1 Degradation of phenols in spiked sewage wastewater using PC-HRP discs.

Time (h)	Concentration of phenolic contaminants (mg/L)	Percentage reduction in phenols
0	30.09 ± 0.18	0
1	5.64 ± 0.03	81
2	2.11 ± 0.11	93
3	1.49 ± 0.27	95

This sample was spiked sewage wastewater. The 0-hour sample is the untreated wastewater. Concentrations of phenols are calculated from a standard graph prepared using the colorimetric method of phenol determination.

Table 2 Physical and chemical analysis of the treated and untreated spiked sewage wastewater.

Sample parameter	Untreated sewage water	PC-HRP treated sewage water
pH	7.0 ± 0.2	7.9 ± 0.1
TDS (mg L^{-1})	880.5 ± 25	665.6 ± 12
Conductivity (µS cm^{-1})	2207 ± 110	1667 ± 70
Salinity (psu)	1.13 ± 0.08	0.84 ± 0.04
Ammonia (ppm)	27.2 ± 0.7	12.0 ± 0.8
Na^+/ Cl^- (mg/L; mV)	238 ± 20; 15.6 ± 1.3	166 ± 20; 6.4 ± 1.0
Nitrite, NO_2^- (mg/L)	1.08 ± 0.1	0.85 ± 0.1
Nitrate, NO_3^- (ng/L)	0.16 ± 0.03	0.07 ± 0.02
BOD (mg/L)	208 ± 26	26 ± 6
COD (mg/L)	620 ± 45	140 ± 18

This sample was spiked sewage wastewater (30 mg/L phenols).

was reflected by the changes in the physical (optical density, pH, TDS, conductivity, salinity) and chemical (ammonia, nitrates, nitrites, BOD, COD) properties of the water (Table 2). The increase in pH or the amount of negative hydroxide ions results in bond formation with the positively charged heavy metal ions present in the wastewater. This creates a dense, insoluble precipitate of metal particles that settle out of wastewater after a given time. The precipitate entraps other particulate matter and salts, resulting in reduced TDS content, conductivity and salinity of wastewater upon treatment. The reduction in the content of ammonia in wastewater after 3 h of treatment might have resulted from its escape during continuous shaking of the wastewater.

Fertilizers and cattle manure are significant sources of nitrate contamination in wastewater. The measured values of nitrate and nitrite in the collected wastewater were calculated to 0.16 mg/L and 1.08 mg/L, respectively. No significant change in the nitrate/nitrite concentration was observed upon biological treatment using HRP; nitrates and nitrites are mainly removed from wastewater using anaerobic bacteria like Pseudomonas species, which convert them to elemental nitrogen that is liberated to the atmosphere.

BOD is a measure of the organic pollution of water. It determines the amount of dissolved oxygen (mg of O_2 consumed per liter of sample) needed by aerobic biological organisms to break down organic material present in a given water sample at certain temperature (e.g. 27 °C) over a specific time period (e.g. 3 days). Unlike BOD, the COD measures the oxygen required to oxidize all carbon compounds in a solution. BOD and COD values of phenol- spiked wastewater were measured before and after treatment. As shown in Table 2, the BOD value dropped from 208 mg/L to approximately 26 mg/L after enzymatic treatment, suggesting the 87% reduction in oxygen demand for biodegradation of organic pollutants. On the other hand, COD reduction of almost 77% was achieved upon treatment with immobilized HRP. Such observations are also reported by other authors [47, 48].

The significant decrease in BOD and COD values of treated spiked-wastewater demonstrated the effectiveness of the proposed method to transform phenolic into precipitable simpler compounds.

3.5 Reusability and Storage Stability of Immobilized HRP

The reusability of immobilized HRP in degradation of the phenolic compounds in sewage wastewater was studied using inverted mixture as described in section 2.9. As we observed 93% degradation in phenolic content in 2 h treatment, we therefore selected cycles of 2 h to investigate the efficiency of phenol removal by immobilized HRP. We found that after 4 cycles of phenol degradation, the efficiency of HRP get reduced to 60%. However, this can be improved by quick washing of HRP-immobilized disc with MilliQ or respective buffer in order to remove the deposited phenolic precipitate and other contaminant which may interfere in enzyme-substrate interactions [12].

Furthermore, storage stability of immobilized HRP was investigated by storing the immobilized preparation of HRP in a sealed polythene bag in refrigerator at 4 °C for two weeks and thereafter studying the phenolic degradation. We found the immobilized HRP could fully retain its activity upon two week storages. This suggests that the immobilized enzyme could be stored at 4 °C for at least two weeks without loss of activity.

4. Conclusions

Direct discharge of industrial effluents enriched in toxic and carcinogenic phenolic pollutants to water bodies can pose a serious threat to human lives. This study investigates the effectiveness of immobilized HRP for phenol degradation in sewage wastewater. The present method reduced the total phenol concentration in sewage wastewater by 91-95%. Beside dephenolization, the treated water showed

overall improvement in quality as reflected by the 77% and 87% reduction of the COD and the BOD values, increased pH and reduced turbidity, dissolved salts concentration and nitrate and nitrite components. Phenols along with other biodegradable compounds were removed as precipitate from wastewater. PEG was used as a protective additive, which helped in retaining enzymatic activity of HRP by 2-3 folds. The results showed that the present approach could be effectively employed for the enzymatic treatment of phenols from industrial wastewater, making the treated water usable for agriculture and non-drinking anthropological activities.

Acknowledgments

The authors thank the Council of Scientific and Industrial Research grant CARDIOMED (BSC0122) and SERB- Department of Science and Technology, Government of India, for providing the financial support to this work. R. A. thanks the Council of Scientific and Industrial Research, Government of India for the award of a senior research fellowship. S. K. thanks SERB-DST, Government of India for the award of Fast Track Young Scientist.

References

[1] Liu, J. Z., Song, H. Y., Weng, L. P., and Ji, L. N. 2002. "Increased Thermostability and Phenol Removal Efficiency by Chemical Modified Horseradish Peroxidase." *J. Mol. Catal., B Enzym.* 18: 225-32.

[2] Asad, S., Dabirmanesh, B., and Khajeh, K. 2014. "Phenol Removal from Refinery Wastewater by Mutant Recombinant Horseradish Peroxidase." *Biotechnol. Appl. Biochem.* 61: 226-9.

[3] Kurnik, K., Treder, K., Skorupa-Klaput, M., Tretyn, A., and Tyburski, J. 2015. "Removal of Phenol from Synthetic and Industrial Wastewater by Potato Pulp Peroxidases." *Water Air Soil Pollut. 226*: 254.

[4] Ikehata, K., Buchanan, I. D., and Smith, D. W. 2003. "Treatment of Oil Refinery Wastewater Using Crude Coprinus Cinereus Peroxidase and Hydrogen Peroxide." *J. Environ. Eng. Sci.* 2: 463-72.

[5] Baker, E. L., Landrigan, P. J., Bertozzi, P. E., Field, P. H., Basteyns, B. J., and Skinner, H. G. 1978. "Phenol Poisoning due to Contaminated Drinking Water." *Arch. Environ. Health* 33: 89-94.

[6] Cohen-Yaniv, V., and Dosoretz, C. G. 2009. "Fate of Horseradish Peroxidase during Oxidation of Monobrominated Phenols." *J. Chem. Technol. Biotechnol.* 84: 1559-66.

[7] Xu, Q., Li, X., Zhou, Y. E., Wei, H., Hu, X. Y., Wang, Y., and Yang, Z. 2012. "An Enzymatic Amplified System for the Detection of 2, 4-Dichlorophenol Based on Graphene Membrane Modified Electrode." *Anal. Methods* 4: 3429-35.

[8] Villegas, L. G. C., Mashhadi, N., Chen, M., Mukherjee, D., Taylor, K. E., and Biswas, N. 2016. "A Short Review of Techniques for Phenol Removal from Wastewater." *Curr. Pollution Rep.* 2: 157-67.

[9] Mohammadi, S., Kargari, A., Sanaeepur, H., Abbassian, K., Najafi, A., and Mofarrah, E. 2015. "Phenol Removal from Industrial Wastewaters: A Short Review." *Desalination and Water Treatment* 53: 2215-34.

[10] Kulkarni, S. J., and Kaware, D. J. P. 2013. "Review on Research for Removal of Phenol from Wastewater." *International Journal of Scientific and Research Publications* 3: 1-4.

[11] Bayramoglu, G., Akbulut, A., and Arica, M. Y. 2013. "Immobilization of Tyrosinase on Modified Diatom Biosilica: Enzymatic Removal of Phenolic Compounds from Aqueous Solution." *J. Hazard. Mater.* 244-245 (2): 528-36.

[12] Cheng, J., Ming Yu, S., and Zuo, P. 2006. "Horseradish Peroxidase Immobilized on Aluminium-Pillared Inter-layered Clay for the Catalytic Oxidation of Phenolic Wastewater." *Water Res.* 40: 283-90.

[13] Jiang, Y., Tang, W., Gao, J., Zhou, L., and He, Y. 2014. "Immobilization of Horseradish Peroxidase in Phospholipid-Templated Titania and Its Applications in Phenolic Compounds and Dye Removal." *Enzyme Microb. Technol.* 55: 1-6.

[14] Cheng, W., and Harper Jr., W. F. 2012. "Chemical Kinetics and Interactions Involved in Horseradish Peroxidase-Mediated Oxidative Polymerization of Phenolic Compounds." *Enzyme Microb. Technol.* 50: 204-8.

[15] Li, H., Li, Y., Cao, H., Li, X., and Zhang, Y. 2011. "Degradation of Pentachlorophenol by a Novel Peroxidase-Catalyzed Process in the Presence of Reduced Nicotinamide Adenine Dinucleotide." *Chemosphere* 83: 124-30.

[16] Nazari, K., Esmaeili, N., Mahmoudi, A., Rahimi, H., and Moosavi-Movahedi, A. A. 2007. "Peroxidative Phenol Removal from Aqueous Solutions Using Activated Peroxidase Biocatalyst." *Enzyme Microb. Technol.* 41: 226-33.

[17] Pan, C., Ding, R., Dong, L., Wang, J., and Hu, Y. 2015. "Horseradish Peroxidase-Carrying Electrospun Nonwoven Fabrics for the Treatment of

o-Methoxyphenol." *J. Nanomaterials* 2015: 1-6.

[18] Zhang, L., Zhao, W., Chen, H., and Cui, Y. 2013. "Enzymatic Synthesis of Phenol Polymer and Its Functionalization." *J. Mol. Catal., B Enzym.* 87: 30-6.

[19] Stanisavljevic, M., and Nedic, L. 2004. "Removal of Phenol from Industrial Wastewaters by Rseradish (Cochlearia Armoracia l) Peroxidase." *Working and Living Environmental Protection* 2: 345-9.

[20] Spahn, C., and Minteer, S. D. 2008. "Enzyme Immobilization in Biotechnology." *Recent Patents on Engineering* 2: 195-200.

[21] Prodanovic, O., Prokopijevic, M., Spasojevic, D., Stojanovic, Z., Radotic, K., Knezevic-Jugovic, Z. D., and Prodanovic, R. 2012. "Improved Covalent Immobilization of Horseradish Peroxidase on Macroporous Glycidyl Methacrylate-based Copolymers." *Appl. Biochem. Biotechnol.* 168: 1288-301.

[22] Kondyurin, A. V., Naseri, P., Tilley, J. M., Nosworthy, N. J., Bilek, M. M., and McKenzie, D. R. 2012. "Mechanisms for Covalent Immobilization of Horseradish Peroxidase on Ion-Beam-Treated Polyethylene." *Scientifica* 2012: 1-28.

[23] Cao, X., Yu, J., Zhang, Z., and Liu, S. 2012. "Bioactivity of Horseradish Peroxidase Entrapped in Silica Nanospheres." *Biosens. Bioelectron.* 35: 101-7.

[24] Lei, C. X., Hu, S. Q., Shen, G. L., and Yu, R. Q. 2003. "Immobilization of Horseradish Peroxidase to a Nano-Au Monolayer Modified Chitosan-Entrapped Carbon Paste Electrode for the Detection of Hydrogen Peroxide." *Talanta* 59: 981-8.

[25] Lin, T. Y., Wu, C. H., and Brennan, J. D. 2007. "Entrapment of Horseradish Peroxidase in Sugar-Modified Silica Monoliths: Toward the Development of a Biocatalytic Sensor." *Biosens. Bioelectron.* 22: 1861-7.

[26] Kumar, S., Kumar, D., Ahirwar, R., and Nahar, P. 2016. "Exploring the Flexible Chemistry of 4-Fluoro-3-Nitrophenyl Azide for Biomolecule Immobilization and Bioconjugation." *Anal. Bioanal. Chem.* 408: 6945-56.

[27] Shannon, L. M., Kay, E., and Lew, J. Y. 1966. "Peroxidase Isozymes from Horseradish Roots. I. Isolation and Physical Properties." *J. Biol. Chem.* 241 (9): 2166-72.

[28] Worthington, K., and Worthington, V. 2011. *Worthington Enzyme Manual.* Worthington: Biochemical Corporation.

[29] Buchanan, I. D., and Nicell, J. A. 1997. "Model Development for Horseradish Peroxidase Catalyzed Removal of Aqueous Phenol." *Biotechnol. Bioeng.* 54: 251-61.

[30] Nahar, P., Wali, N. M., and Gandhi, R. P. 2001. "Light-induced Activation of an Inert Surface for Covalent Immobilization of a Protein Ligand." *Anal. Biochem.* 294: 148-53.

[31] Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. 1985. "Measurement of Protein Using Bicinchoninic Acid." *Anal. Biochem.* 150: 76-85.

[32] Wu, Y., Taylor, K. E., Biswas, N., and Bewtra, J. K. 1998. "A Model for the Protective Effect of Additives on the Activity of Horseradish Peroxidase in the Removal of Phenol." *Enzyme Microb. Technol.* 22: 315-22.

[33] Zatar, N. A., Abu-Eid, M. A., and Eid, A. F. 1999. "Spectrophotometric Determination of Nitrite and Nitrate Using Phosphomolybdenum Blue Complex." *Talanta* 50: 819-26.

[34] Central Water Commission, G.o.I. 1999. "Standard Analytical Procedures for Water Analysis." CSMRS Building, 4th Floor, Olof Palme Marg, Hauz Khas, New Delhi-110016 (INDIA).

[35] Davies, H. M., and Manning, J. R. 2008. "Catalytic C-H Functionalization by Metal Carbenoid and Nitrenoid Insertion." *Nature* 451: 417-24.

[36] Bora, U., Sharma, P., Kumar, S., Kannan, K., and Nahar, P. 2006. "Photochemical Activation of a Polycarbonate Surface for Covalent Immobilization of a Protein Ligand." *Talanta* 70: 624-9.

[37] Prodanovic, O., Prokopijevic, M., Spasojevic, D., Stojanovic, Z., Radotic, K., Knezevic-Jugovic, Z. D., and Prodanovic, R. 2012. "Improved Covalent Immobilization of Horseradish Peroxidase on Macroporous Glycidyl Methacrylate-based Copolymers." *Appl. Biochem. Biotechnol.* 168: 1288-301.

[38] Masuda, Y., Kugimiya, S. I., and Kato, K. 2014. "Improvement of Thermal-Stability of Enzyme Immobilized onto Mesoporous Zirconia." *Journal of Asian Ceramic Societies* 2: 11-9.

[39] Temocin, Z., and Yigitoglu, M. 2009. "Studies on the Activity and Stability of Immobilized Horseradish Peroxidase on Poly(Ethylene Terephthalate) Grafted Acrylamide Fiber." *Bioprocess Biosyst. Eng.* 32: 467-74.

[40] Ai, J., Zhang, W., Liao, G., Xia, H., and Wang, D. 2016. "Immobilization of Horseradish Peroxidase Enzymes on Hydrous-Titanium and Application for Phenol Removal." *RSC Advances* 6: 38117-23.

[41] Mukhopadhyay, A., Dasgupta, A. K., and Chakrabarti, K. 2013. "Thermostability, pH Stability and Dye Degrading Activity of a Bacterial Laccase are Enhanced in the Presence of Cu_2O Nanoparticles." *Bioresource Technology* 127: 25-36.

[42] Klibanov, A. M., Tu, T. M., and Scott, K. P. 1983. "Peroxidase-Catalyzed Removal of Phenols from Coal-Conversion Waste Waters." *Science (New York,*

N.Y.) 221: 259-61.

[43] Nicell, J. A., Saadi, K. W., and Buchanan, I. D. 1995. "Phenol Polymerization and Precipitation by Horseradish Peroxidase Enzyme and an Additive." *Bioresource Technology* 54: 5-16.

[44] Nakamoto, S., and Machida, N. 1992. "Phenol Removal from Aqueous Solutions by Peroxidase-Catalyzed Reaction Using Additives." *Water Research* 26: 49-54.

[45] Mao, L., Luo, S., Huang, Q., and Lu, J. 2013. "Horseradish Peroxidase Inactivation: Heme Destruction and Influence of Polyethylene Glycol." *Scientific Reports* 3: 3126.

[46] Tatsumi, K., Wada, S., and Ichikawa, H. 1996. "Removal of Chlorophenols from Wastewater by Immobilized Horseradish Peroxidase." *Biotechnol. Bioeng.* 51: 126-30.

[47] Wagner, M., and Nicell, J. A. 2001. "Peroxidase-Catalyzed Removal of Phenols from a Petroleum Refinery Wastewater." *Water Science and Technology: A Journal of the International Association on Water Pollution Research* 43: 253-60.

[48] Bilal, M., and Asgher, M. 2015. "Dye Decolorization and Detoxification Potential of Ca-Alginate Beads Immobilized Manganese Peroxidase." *BMC Biotechnol.* 15: 111.

View publication stats

Available online at www.ijapie.org

INTERNATIONAL JOURNAL OF ADVANCED PRODUCTION AND INDUSTRIAL ENGINEERING

IJAPIE-2017-02-211, Vol 2 (2), 01-04

IJAPIE
Connecting Science & Technology with Management.
A Journal for all Products & Processes.

An Analysis of Selected Parameters in an Energy Efficient Sustainable Manufacturing System Using AHP

Subrata Kumar Patra[1*], Tilak Raj[2], B.B. Arora[3]
([1]*Department of Production Engineering, G.B. Pant Institute of Technology, New Delhi, India*
[2]*Department of Mechanical Engineering, YMCAUST, Faridabad, India*
[3]*Department of Mechanical Engineering, Delhi Technological University, Delhi, India*
**Email: patrask2005@yahoo.co.in*

Abstract : Sustainability has taken a pivotal role to enhance the business opportunities by means of producing eco-friendly products using sustainable manufacturing practices. In order to survive in the global competitive market, the need of the hour is to focus on sustainable manufacturing practices. This is furthermore required due to scarcity of natural resources, increased cost of labor/raw material/ machinery/energy, stringent government norms/regulations, higher aspirations of consumers and the like. Increased global competition is forcing the manufacturers to shift towards eco-friendly products - sustainable manufacturing is the pathway in this direction.
This paper will aim to analyze some selected parameters in an energy efficient Sustainable Manufacturing system. In the present work Analytical Hierarchy Process (AHP) has been used as a decision making technique.

Keywords: AHP - Analytical Hierarchy Process, MADM- Multiple Attribute Decision Making, CI- Consistency Index, CR- Consistency Ratio, RI- Random Index

1. INTRODUCTION

Any manufacturing activity is an assortment of men, machine, material, money, method and management. These when coupled together can give rise to the desired output in any manufacturing environment. To carry out any manufacturing activity raw materials are essential in some form or the other. These are extracted from the nature and are then transformed into various kinds of products (semi-finished type or finished type) for end users with the help of appropriate machines/tools/process in a suitable environment by involving workmen. With the declining stock of raw materials and resources and the rising business competition, there is an urgent need for improving the manufacturing effectiveness. As per Subrata Kumar Patra et al. **[2015]** apart from depletion of natural resources various allied problems for the future will be scarcity of water especially potable water, fossil fuel, energy, land filling sites in one hand and increase of pollution level, global warming, chronic diseases etc. on the other. The role of energy in any manufacturing activity is inevitable. There is an urgent need for the manufacturing sector to work towards converting traditional manufacturing system into an energy efficient sustainable manufacturing system.

2. SUSTAINABLE MANUFACTURING

A Traditional manufacturing system eternally focuses for enhancing the quality of manufacture as well as increasing the productivity at the lowest possible cost. Traditional manufacturing system generally has very little focus on conserving natural resources. The environmental damage caused due to the Traditional manufacturing process acts as a barrier towards Sustainable development. On the other hand, Sustainable manufacturing system aims towards manufacturing by conserving natural resources and continually strives towards reducing environmental pollution, waste and emission by adopting efficient environmental friendly practices.

Study of literatures revealed various issues related to Energy efficient Sustainable manufacturing system. These may be broadly categorized as:

- Environmental issues
- Sustainable manufacturing issues
- Energy issues

Further studies helped to identify sub-level elements out of which the following are selected as important parameters for the present work. These are listed in Table 1.1.

Table 1.1: Selected elements in an Energy efficient sustainable manufacturing system

P1	Waste management		P6	Technology issues
P2	Pollution /Emission Control		P7	Manufacturing Process
P3	Environment consciousness		P8	Utilization of Energy
P4	Resource Conservation		P9	Energy Efficiency
P5	Recycling Practices		P10	Use of Alternate Energy

Fig.1.1 depicts the selected issues and the selected parameters in an energy efficient Sustainable Manufacturing System.

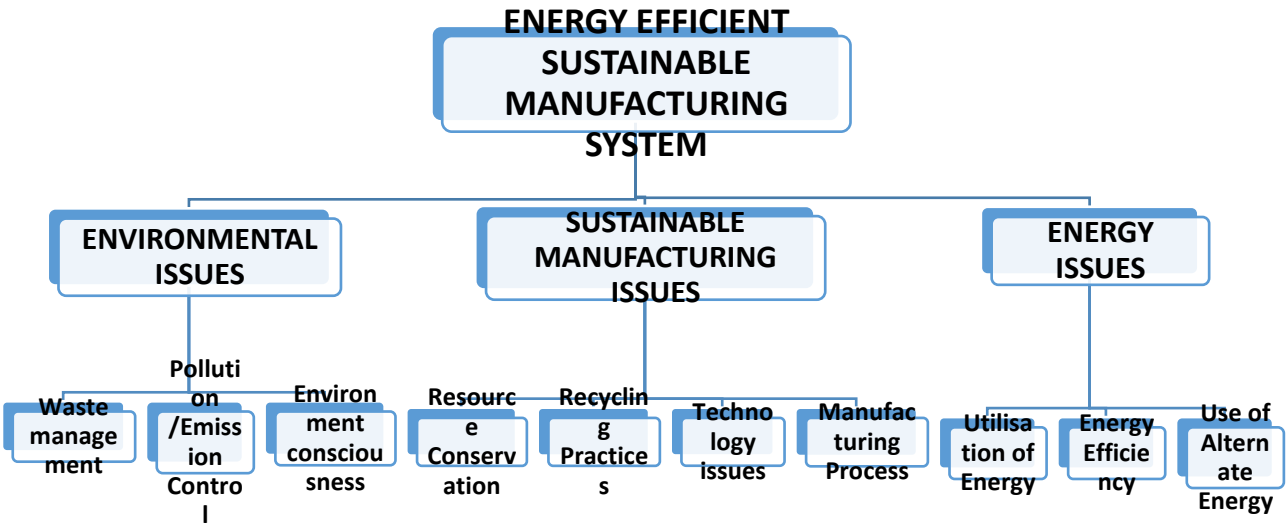


Fig.1.1Various parameters in an energy efficient Sustainable Manufacturing System

3. METHODOLOGY

Any Manufacturing system uses raw materials that are converted into useful products using suitable machines and processes. Various associated parameters are changed from time to time depending on the product and process variability in order to optimize the system. Decision making under such a complex dynamic situation based on multiple alternatives is a challenging task for the engineers, planners, designers, decision makers and the management. Decision making may be done by individual judgment or by a group of experts. In case of subjective opinions decision makers are required to give some subjective values (based on their judgmental power) to quantify them.

Various types of Multiple Attribute Decision Making (MADM) techniques are used to select the appropriate alternatives from a large number of available choices. In the present work Analytical Hierarchy Process (AHP) has been taken up as a decision making tool because of its wider acceptability, adaptability with various situations, easiness to use and the like. Moreover AHP method is capable of solving both objective as well as subjective attributes. Ranking of the alternatives are done using Pair wise comparison of attributes followed by a scale of relative importance. A Questionnaire based survey is used to find the relative importance of these selected parameters (Table 1.1).

The step by step criteria followed in the AHP analysis are:

a. Developing questionnaire for Pair wise comparison
b. Formation of Comparison matrix
c. Normalization
d. Evaluation and Consistency analysis

The scale for comparison used in this paper is given below *(Saaty & Vargas, 1991).*

1- Equal importance of both factors
3-Moderate importance of one factor over another
5-Strong or essential importance
7- Very strong importance
9- Extreme importance.

Integer values of 2, 4, 6, and 8 are used for intermediate values. The scale has been used with the assumption that out of two attributes say P1 and P2, if attribute P1 has very strong importance than P2 and say it is 7, then P2 must be absolutely less important than P1and is valued as 1/7. The matrix formation has been done by pair wise comparison followed by normalizing the matrix and Consistency analysis. Consistency analysis is done by first calculating the consistency measure followed by calculating Consistency Index (CI) and Consistency Ratio (CR). CR is a ratio of CI

and RI where RI is called Random Index and is obtained from RI index chart.

4. RESULTS AND DISCUSSION:

Table 1.2 depicts the calculation chart based on AHP method.

Table 1.2: Calculation chart used in AHP

i

Parameters	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
P1	1.000	0.500	6.000	2.000	4.000	1.000	3.000	0.333	1.000	5.000
P2	2.000	1.000	8.000	4.000	6.000	1.000	5.000	1.000	3.000	7.000
P3	0.167	0.125	1.000	0.250	0.500	0.143	0.333	0.111	1.000	1.000
P4	0.500	0.250	4.000	1.000	2.000	1.000	1.000	1.000	0.200	1.000
P5	0.250	0.167	2.000	0.500	1.000	0.200	1.000	0.143	1.000	1.000
P6	1.000	1.000	7.000	1.000	5.000	1.000	4.000	1.000	2.000	6.000
P7	0.333	0.200	3.000	1.000	1.000	0.250	1.000	0.167	0.500	2.000
P8	3.000	1.000	9.000	1.000	7.000	1.000	6.000	1.000	4.000	8.000
P9	1.000	0.333	1.000	5.000	1.000	0.500	2.000	0.250	1.000	4.000
P10	0.200	0.143	1.000	1.000	1.000	0.167	0.500	0.125	0.250	1.000
TOTAL	9.450	4.718	42.000	16.750	28.500	6.260	23.833	5.129	13.950	36.000

P1	Waste management
P2	Pollution /Emission Control
P3	Environment consciousness
P4	Resource Conservation
P5	Recycling Practices
P6	Technology issues
P7	Manufacturing Process
P8	Utilisation of Energy
P9	Energy Efficiency
P10	Use of Alternate Energy

ii

Parameters	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	TOTAL	AVERAGE	Consistency Measure
P1	0.106	0.106	0.143	0.119	0.140	0.160	0.126	0.065	0.072	0.139	1.176	*0.118*	10.868
P2	0.212	0.212	0.190	0.239	0.211	0.160	0.210	0.195	0.215	0.194	2.037	*0.204*	11.070
P3	0.018	0.026	0.024	0.015	0.018	0.023	0.014	0.022	0.072	0.028	0.258	*0.026*	11.296
P4	0.053	0.053	0.095	0.060	0.070	0.160	0.042	0.195	0.014	0.028	0.770	*0.077*	10.771
P5	0.026	0.035	0.048	0.030	0.035	0.032	0.042	0.028	0.072	0.028	0.376	*0.038*	11.175
P6	0.106	0.212	0.167	0.060	0.175	0.160	0.168	0.195	0.143	0.167	1.552	*0.155*	10.804
P7	0.035	0.042	0.071	0.060	0.035	0.040	0.042	0.032	0.036	0.056	0.450	*0.045*	11.026
P8	0.317	0.212	0.214	0.060	0.246	0.160	0.252	0.195	0.287	0.222	2.164	*0.216*	10.959
P9	0.106	0.071	0.024	0.299	0.035	0.080	0.084	0.049	0.072	0.111	0.929	*0.093*	11.445
P10	0.021	0.030	0.024	0.060	0.035	0.027	0.021	0.024	0.018	0.028	0.288	*0.029*	11.136
												CI	0.117
												RI	1.490
												C. RATIO	0.079

The result gives a CR ratio of 0.079 which is well below the limit of 0.1 and is thus acceptable. This implies that the judgments are quite consistent and the overall work is meaningful. The consistency measure gives the relative ranks of these 10 parameters and is given below:

PARAMETER NO.	PARAMETER NAME	RANK
P9	Energy Efficiency	I
P3	Environment consciousness	II
P5	Recycling Practices	III
P10	Use of Alternate Energy	IV
P2	Pollution /Emission Control	V
P7	Manufacturing Process	VI
P8	Utilization of Energy	VII
P1	Waste management	VIII
P6	Technology issues	IX
P4	Resource Conservation	X

This implies that energy efficiency is the most important parameter in an energy efficient sustainable manufacturing system. This is followed by environmental consciousness and recycling practices and so on.

5. CONCLUSION

The objective of the present work was to find out the relative ranking/ importance of the selected parameters in a Sustainable manufacturing environment using AHP method. Other methods that could have been used are TOPSIS, Artificial Neural Network, Genetic Algorithm etc. In the present study AHP method has been used to rank these alternatives. Based on the consistency measures the ranking is P9- P3- P5- P10- P2- P7- P8- P1- P6- P4. This means that the energy efficiency is the most important factor in any sustainable manufacturing system. The manufacturers are therefore expected to switch over to more energy saving appliances, focus on research and development for energy efficient machines and adopt energy efficient processes for the development that will be truly sustainable.

REFERENCES

1. A.K. Kulatunga, P. R. Jayatilaka, M. Jayawickrama, “Drivers and barriers to implement sustainable manufacturing concepts in Sri Lankan manufacturing sector”, Department for Production Engineering, University of Peradeniya, 2013, Sri Lanka
2. Anja Kollmuss, Julian Agyeman, “Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior?”, Environmental Education Research, Vol. 8, No. 3, 2002, pp. 239-260
3. Askiner Gungor, Surendra M. Gupta “Issues in environmentally conscious manufacturing and product recovery: A survey”, Computers & Industrial Engineering, Vol. 36, 1999, pp. 811-853

4. Barbara Linke1, Yu-Chu Huang, David Dornfeld, "Establishing Greener Products and Manufacturing Processes", International Journal of Precision Engineering and Manufacturing, Vol. 13, No. 7, 2012, pp.1029-1036
5. Claudia Ghisetti , Massimiliano Mazzanti , Susanna Mancinelli , Mariangela Zoli, "Do financial constraints make the environment worse off? Understanding the effects of financial barriers on environmental innovations", SEEDS Working Paper, Vol. 1, 2015, pp. 1-21
6. E. B. Brousseau & S. S. Dimov & D. T. Pham (2010), "Some recent advances in multi-material micro- and nanomanufacturing," Int J Adv Manuf Technol, Vol.47, pp.161–180
7. Geir B. Asheim, "Sustainability", The World Bank, 1994
8. Geoff Miller, Janice Pawloski, Charles Standridge, "A case study of lean, sustainable manufacturing", Journal Of Industrial Engineering and Management, Vol. 3, No.1, 2010, pp. 11-32
9. Hameed O Adebambo, Hasbullah Ashari, Norani Nordin, "Driving Factors of Sustainable Environmental Manufacturing Practices in Malaysia", Journal of Technology and Operations Management, Vol. 9, Issue 2, 2014, pp. 7-16
10. Hartmut Kaebernick, Sami Kara, "Environmentally Sustainable Manufacturing: A Survey on Industry Practices", 13th CIRP International Conference on Life Cycle Engineering, 2006
11. J. Kopac, "Achievements of sustainable manufacturing by machining", Journal of Achievements in Materials and Manufacturing Engineering, Vol. 34, Issue 2, 2009, pp. 180-187
12. Ming Chen, "Sustainable Recycling of Automotive Products in China: Technology and Regulation -Key Manufacturing Technology & Equipment for Energy Saving and Emissions Reduction in Mechanical Equipment Industry", JOM: The Journal of The Minerals, Metals & Materials Society, Vol. 58, Issue 8, 2006, pp. 23-26
13. Pankaj R. Shelar, Ramesh R. Lekurwale, " Selecting Appropriate Cutting Tool Insert For Turning Using Analytical Hierarchy Process and Weighted Product Method", Proceedings of IRF-ieeeforum International Conference, ISBN: 978-93-85973-95-6, 2016, pp. 31-34
14. Paul R. Kleindorfer, Kalyan Singhal, Luk N. Van Wassenhove, "Sustainable Operations Management", Production and Operations Management, Vol. 14, No. 4, 2005, pp. 482–492
15. R. W. Saaty, "The Analytic Hierarchy Process- What It Is and how It Is used", Mathl Modelling, Vol. 9, No. 3-5, 1987, pp. 161-176
16. Roger C. Herdman, "Perspectives on the Role of Science and Technology in Sustainable Development", OTA-ENV-609, 1994
17. Saaty T L (1980),The Analytic Hierarchy Process, McGraw Hill International
18. Saaty T L, " The Analytic Hierarchy Process", McGraw Hill International, 1980
19. Shaw C. Feng and Che B. Joung, "An Overview of a Proposed Measurement Infrastructure for Sustainable Manufacturing", Proceedings of the 7th Global Conference on Sustainable Manufacturing, 2009
20. Shaw C. Feng, Che B. Joung, "An Overview of a Proposed Measurement Infrastructure for Sustainable Manufacturing", Proceedings of the 7th Global Conference on Sustainable Manufacturing, 2009
21. Subrata Kumar Patra, Tilak Raj, B.B. Arora, "Sustainability Issues In Energy Efficient Manufacturing Systems - A Review", International Journal of Engineering and Manufacturing Science, ISSN 2249-3115, Vol. 5, No. 1, 2015, pp. 1-8
22. Thomas L. Saaty, "Decision making with the analytic hierarchy process", Int. J. Services Science, Vol. 1, No. 1, 2008 , pp. 83-98
23. United Nations Department of Economic and Social Affairs, "Sustainable Consumption and Production: Energy and Industry", May 2006
24. Zhongde Shan, Shaoyan Qin, Qian Liu, Feng Liu, " Key manufacturing technology & equipment for energy saving and emissions reduction in mechanical equipment industry", International Journal of Precision Engineering and Manufacturing, Vol. 13, No. 7, 2012, pp. 1095-1100.

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Analytical modeling to study the effect of hydrogen plasma on the growth of multi-walled carbon nanotubes

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 J. Phys.: Conf. Ser. 836 012017

(http://iopscience.iop.org/1742-6596/836/1/012017)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 14.139.251.107
This content was downloaded on 06/06/2017 at 11:16

Please note that terms and conditions apply.

You may also be interested in:

New practical method of homogeneous dispersion of multi-walled carbon nanotubes (MWCNTs) into Mg matrix composites
Teng Zhang, Shuangming Du, Wanchang Sun et al.

Multi-Walled Carbon Nanotubes Growing Vertically from Root Particles
Hui Lin Chang and Cheng Tzu Kuo

An Experimental Study on Buoyancy Induced Convective Heat Transfer in a Square Cavity using Multi-Walled Carbon Nanotube (MWCNT)/Water Nanofluid
Pranit Joshi and Arvind Pattamatta

Electrical, dielectric and surface wetting properties of multi-walled carbon nanotubes/nylon-6 nanocomposites
Long Yun-Ze, Li Meng-Meng, Sui Wan-Mei et al.

Growth Valley Dividing Single- and Multi-Walled Carbon Nanotubes: Combinatorial Study of Nominal Thickness of Co Catalyst
Kazunori Kakehi, Suguru Noda, Shigeo Maruyama et al.

Etching of Multi-Walled Carbon Nanotubes Using Energetic Plasma Ions
Jung-Hyun Cho and Gon-Ho Kim

Effects of different processing techniques on multi-walled carbon nanotubes/silicone rubber nanocomposite on tensile strength properties
N Mazlan, M Jaafar, A Aziz et al.

Influence of Temperature on the Conductivity of Multi-walled Carbon Nanotube Interconnects
Lu Qi-Jun, Zhu Zhang-Ming, Yang Yin-Tang et al.

Analytical modeling to study the effect of hydrogen plasma on the growth of multi-walled carbon nanotubes

U Sharma[a], N Gupta, R Gupta[1]and S C Sharma

Department of Applied Physics, Delhi Technological University, Delhi-110042, India

Email: [a)]umangsharma2810@gmail.com

Abstract. An analytical model has been developed to study the effect of hydrogen plasma on the growth of Multi-walled Carbon Nanotubes (MWCNTs). The model incorporates the charging rate of the MWCNT, kinetics of plasma species, namely, electrons, positively charged ions, and neutral atoms, growth rate of MWCNT in the presence of CH_4/H_2 plasma. Numerical calculations have been carried out to study the influence of hydrogen number density on the dimensions of MWCNT for typical glow discharge plasma parameters. The impact of hydrogen on the hydrocarbon density has also been studied in the present investigation. The present work can serve as a major tool in better understanding of highly proficient field emitters based on MWCNTs.

1. Introduction

The carbon nanotube (CNT) is a novel carbon structure found since 1991 [1]. CNTs due to their extraordinary mechanical properties, have been getting much attention for a wide variety of applications. Maruyama et al. [2] prospered in producing high-purity single-walled CNTs (SWCNTs) with a high number density. Hata et al. [3] discovered a way to grow high-purity SWCNTs with a high growth rate by adding a small amount of H_2O vapor to the carbon source. They advocated that OH radicals worked as a weak oxidizer to eradicate impurities without damaging the CNT. Zhang et al. [4] argued the effects of H radicals in the gas phase on SWCNT growth. They suggested that the balance of carbon and hydrogen radicals were vital for CNT growth. Plasma enhanced chemical vapour deposition (PECVD) has appeared as a key growth technique to yield vertically-aligned nanotubes [5]. Researchers using plasma processing have examined the correlation between a plasma gas phase and deposition mechanisms. Ostrikov et al. [6] examined $CH_4/H_2/Ar$ plasma for synthesis of carbon nanostructures. They noted that H atoms affected catalyst activity and that cations were provided as precursors for the carbon nanostructure. Denysenko et al. [7] exhibited that C_xH_y ions increased with the supply of the carbon source and that H atoms also increased with the C_xH_y molecules formed in the $CH_4/H_2/Ar$ plasma. Hydrogen plays a dual role in the PECVD growth of CNT. It acts as an activator for the dissociation of hydrocarbon species into more active species and also as an etching reagent of CNT. In the present work, we analyze the effect of hydrogen on the growth of CNTs in plasma processing, analytically. In section 2, we develop a theoretical model incorporating charging rate of the CNT, kinetics of plasma species i.e., neutral atoms, growth rate of CNT in the presence of CH_4/H_2 plasma. Results and discussions are given in section 3. Finally, the conclusion is given in section 4.

2. Model

We consider plasma comprising electrons, positively charged ions of types A and B, neutral atoms of type A and B, and CNT grown in the existence of plasma. The positively charged ions are anticipated to be singly ionized. For simplicity of the problem, average radius of MWCNT has been considered.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

Charge neutrality equation

$$\frac{dZ}{d\tau} = n_{iAct} + n_{iBct} - \gamma_e n_{ect}$$

where Z=the charge on the CNT, n_{ijct}= the ion collection current for A and B type ions and n_{ct}= the number density of the CNT. Growth rate equation of the mass of CNT

$$\frac{dm_{ct}}{d\tau} = (m_A \gamma_A n_{Act} + m_{iA} \gamma_{iA} n_{iAct})$$

where $m_{ct} = \frac{4}{3}\pi a^3 \rho_{ct}$is the mass of CNT and the terms are the gain in the mass density due to the collection of atomic and ionic species A. Energy balance equation for neutral atoms

$$\begin{aligned}
&\frac{d}{d\tau}\left[\frac{3}{2}(n_A + n_B)k_B T_n\right] \\
&= \left[\left(\frac{3}{2}k_B\right)(\alpha_A n_e n_{iA} + \alpha_B n_e n_{iB})(T_e + T_i)\right) + \left(\alpha_A n_e n_{iA} I_{pA} + \alpha_B n_e n_{iB} I_{pB}\right)\Big] \\
&+ \left(\frac{3}{2}k_B\right)\{n_e(v_{eA}\delta_{eA} + v_{eB}\delta_{eB})(T_e - T_n)[(v_{iAA}\delta_{iAA} + v_{iAB}\delta_{iAB})n_{iA} \\
&+ (v_{iBA}\delta_{iBA} + v_{iBB}\delta_{iBB})n_{iB}](T_i - T_n)\} + [\left(\frac{3}{2}k_B\right)n_{ct}[(1-\gamma_{iA})n_{iAc} + n_{iBc}]T_{ct} \\
&- \left(\frac{3}{2}k_B\right)n_{ct}[n_{Ac}[\gamma_A T_n + \delta_{Act}(1-\gamma_A)(T_n - T_{ct})] + n_{Bc}\delta_{Bct}(T_n - T_{ct})]] \\
&- \left(\frac{3}{2}k_B\right)(\beta_A n_A + \beta_B n_B)T_n - E_{diss}
\end{aligned}$$

where the first term is the power gained per unit volume by the neutral species due to the recombination of electrons and positively charged ions, the second term is the rate of power gained per unit volume by neutral atoms in elastic collisions with electrons and positively charged ions, and the third term is the energy gained per unit volume per second due to the formation of neutrals at the surface of the CNT due to accretion and elastic collisions with CNT. The fourth term refers to the thermal energy lost per unit volume per unit time by neutral atoms due to ionization. The last term is the energy dissipation rate per unit volume by neutral atoms into the surrounding atmosphere.

3. Results and discussions

We have solved the above mentioned equations along with others mentioned in the work of Sharma and Tewari [8] simultaneously for the charging of the CNT and for the kinetics and energy balance of electrons, ions, neutrals with appropriate boundary conditions, viz., at τ=0, n_{ct}=10^6 cm^{-3}, n_{iA0} = 0.58n_{e0}, n_{iB0}= 0.42 n_{e0}, n_{A0} = n_{B0}= 5.1 x 10^{11} cm^{-3}, n_{e0} = 1.1 x 10^{12} cm^{-3}, T_{e0}= 0.6 eV, T_{i0}= 2400 K, and T_{n0}= T_{ct}= 2000 K.

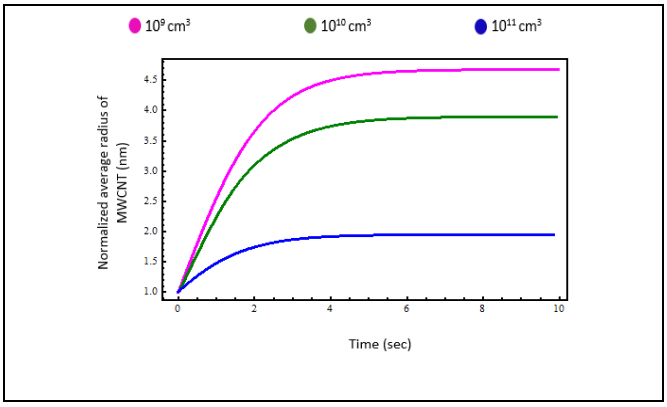


Figure 1. Time variation of normalized average radius of CNTs for different hydrogen ion density

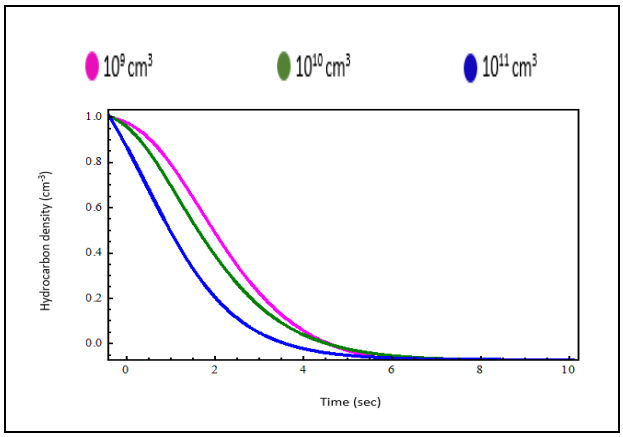


Figure 2. Time variation of hydrocarbon density of MWCNTs for different hydrogen ion density

Figure 1illustrates variation of the normalized radius of a CNT with time for different hydrogen ion number densities. It can be seen that the normalized radius of the CNT first increases with time and then attains a saturation value. The figure also shows the decrease of the normalized radius with the hydrogen ion density. This happens because for larger values of the hydrogen density, the hydrocarbon density decreases. Figure 2 illustrates variation of the hydrocarbon density with time for different hydrogen ion number densities. It can be seen that with the increase in hydrogen ion densities with time, the hydrocarbon density decreases which further affects the radius of MWCNT.

4. Conclusions

An analytical model for the growth kinetics of CNT in plasma has been developed. The approach is based on the charge neutrality and number density of electrons, positively charged ions, neutral atoms, and CNTs. It is analysed from our study that on increasing the hydrogen ion density, the hydrocarbon density and the average radius of MWCNT decreases. The above mentioned results can prove to be beneficial in the fabrication of MWCNTs.

References

[1] Iijima S 1991 *Nature* **354** 56
[2] Maruyama S, Kojima R, Miyauchi Y, Chiashi S and Kohno M 2002 *Chem. Phys. Lett.* **360** 229
[3] Hata K, Futaba DN, Mizuno K, Namai T, Yumura M and Iijima S2004 *Science* **306** 1362
[4] Zhang G, Mann D, Zhang L, Javey A, Li Y, Yenilmez E, Wang Q McVittie J P, Nishi Y, Gibbons J and Dai H 2005 *Proc. Nat. Acad. Sci. USA* **102** 16141
[5] Ostrikov K, Tzakadze Z,Rutkevych P P, Long J D, Xu S and Denysenko I 2005 *Contrib. Plasm. Phys.* **45** 7
[6] Ostrikov K 2005 *Rev. Mod. Phys.* **77** 489
[7] Denysenko I B, Xu S, Long J D, Rutkevych P P, Azarenkov N A and Ostrikov K 2004 *J. App. Phys.* **95** 2713
[8] Sharma S C and Tewari A 2011 *Phys. Plasmas* **18** 063503

Research Article 2017, **2**(6), 378-383 Advanced Materials Proceedings

Anticancer property of green material through computational approach

Nupur Jauhari[1], Gajula MNV Prasad[2], Neelam Sharma[3], Navneeta Bharadvaja[1*]

[1]*Department of Biotechnology, Delhi Technological University, Bawana Road, Delhi-110042 India*
[2]*Institute of Biotechnology, PJTSAU, Rajendra Nagar, Hyderabad-500030, India*
[3]*ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110 012, India*

[*]Corresponding author: E-mail: navneetab@dtu.co.in

Received: 31 March 2016, Revised: 30 September 2016 and Accepted: 18 April 2017

DOI: 10.5185/amp.2017/606
www.vbripress.com/amp

Abstracts

The purpose of research: Colorectal cancer is third most prevalent cancer in developed countries with increasing cases in developing countries. Studies show that the chronic intestinal inflammation is associated with increased risk of developing colorectal cancer. Inflammation is an immunological response to external damaging stimuli and is govern by an endogenous pyrogen and pleiotropic pro-inflammatory cytokine, tumor necrosis factor-alpha (TNF-α). TNF-α plays an important role in the development of humoral immune response. Production of TNF-α has been implicated in various other pathologies including diabetes, osteoporosis, multiple sclerosis and inflammatory bowel diseases also. Several studies have shown that anti-inflammatory effect of stigmasterol, a phytosterol of an endangered medicinal plant *Chlorophytum borivilianum*, is mediated by suppression of TNF-α. The latter is synthesized as a membrane-anchored precursor. The soluble form of TNF-α is released into extracellular space by tumor necrosis factor alpha converting enzyme (TACE), a multidomain metalloproteinase.

Principal results: We have investigated the anti-cancer effect of the green material against colon cancer by using computational molecular docking and molecular dynamics simulations approach. With this study, we tried to explore stigmasterol as a potential inhibitor of TACE. The active cleft of TACE has a catalytic zinc residue at its centerpentacoordinate by the three-imidazole N_2 atoms of His 405, His 409, His 415. Along with these residues Glu 406, Met 345, Pro 437, Trp 312, Asp 344 around the zinc atom constitute the active site. The docked complex formed a hydrogen bond with Gly 346 and showing hydrophobic interactions with other active site residues. Molecular dynamics simulations confirmed that hydrogen-bond connectivity has not been lost throughout our simulations of 30ns. Computed RSMD suggests that docked complex is stable.

Major conclusions: Therapeutic blockade of TACE could be highly beneficial in case of chronic inflammatory conditions including colon cancer. In the present study, we conclude that stigmasterol, a green material may be considered as a possible therapeutic agent in the treatment of colon cancer and the same can certainly be confirmed through *in vivo* studies and clinical trials to confirm their effectiveness in patients.

Keywords: Colon cancer, molecular docking, stigmasterol, TACE, TNF-α.

Introduction

Colorectal cancer is one of the third commonly diagnosed cancers. This non-skin cancer is one of the major causes responsible for death in both men and women in developed countries. According to American Cancer Society, Cancer Facts & Figures (2014), there are 1,36,830 estimated new cases of colon cancer and 50,310 estimated deaths in both the sexes [**1**]. Colon is the part of the digestive system or gastrointestinal system. Most colorectal cancers originate from adenomatous polyp. Polyps are the growth that begins in the inner lining of the colon and moves toward the center. Most of the polyps are noncancerous. Only a few types of polyps can become cancer known as adenomas. Colon cancer grows slowly which takes 10-15 years [**2**]. Risk factors for this disease include physical inactivity, obesity, and high consumption of meat, smoking, and heavy alcohol intake. A family background of colon cancer or adenomatous polyps and personal history of inflammatory bowel disease are the other risk factors associated with colon cancer. Inflammation and tumorogenesis is a grievous combination, well supported by various researchers in last decade.

Copyright © 2017 VBRI Press

Inflammation is a potential source of sporadic and genetic colon cancer.

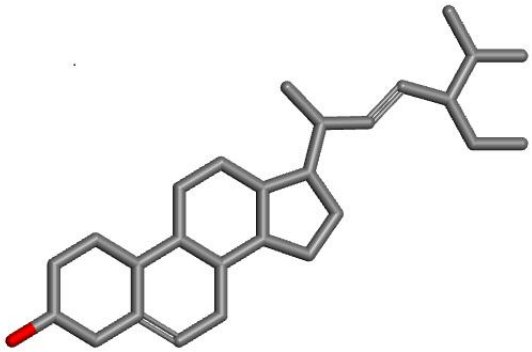


Fig. Structure of Stigmasterol.

Chronic inflammatory bowel disease ulcerative colitis is an important risk factor for the development of colon cancer [**3**]. Treatment of colon cancer with the oral drugs is possible only when diagnosed at the early stage of cancer development. Some FDA approved drugs for colon cancer are Avastin, Bevacizumab, Camptosar, Cetuximab, 5FU (Fluorouracil Injection). Surgery, radiotherapy, and chemotherapy are advised at the latter stage of colon cancer, which are associated with a reoccurrence of the disease and side effects like fatigue, constipation or diarrhea, a temporary or permanent colostomy, bladder irritation etc. Apart from the continuous development of new approaches to combat this baneful disease, not all the approved drugs can be prescribed for continuous long-term use due to drawing some adverse effects. Few advanced imaging techniques enabled the progression of treatment approach in diagnosis of hepatic metastatic disease and rectal cancer with the limitations like technical, economical and logistic challenges [**4**]. Besides the advancement of therapeutic strategies; treatment of colorectal cancer need to be more specific and effective [**5**]. On the contrary, medicinal plants or green material has a great potential to develop herbal drugs with very low or no adverse effects.

'Medicinal plants as Green material' is now becoming a major area of interest to the scientific society across the globe for drug development. Medicinal plants have the capacity to lower down the effects of cholesterol due to the presence of phytosterols. The most important phytosterols are sitosterol, campesterol, and stigmasterol. *Chlorophytum borivilianum*, 'Safed Musli', is one of the important medicinal plants having stigmasterol as major secondary metabolite with high medicinal value. *C. borivilianum*, also known as "the golden root", is an endangered herb that belongs to family 'Liliaceae'. Tuberous roots of this herb contain immunomodulatory and adaptogenic properties and are used to treat impotency, sterility and enhance male potency [**6**]. It is being used to cure physical illness and weakness, diabetes, arthritis, natal and postnatal problems, rheumatism and joint pains. This herb is an antimicrobial, anti-inflammatory and antitumor agent, also used to medicate diarrhea, dysentery, gonorrhea and leucorrhea. Medicinal properties and various biological activities viz. aphrodisiac, anti-oxidant, anti-cancer and immune booster are attributed by its major metabolite stigmasterol. Experimental studies suggest that due to having an anti-inflammatory effect, stigmasterol may offer protection from the most common cancers such as colon, breast, and prostate cancer. The chemopreventive effect of stigmasterol on *in vivo* cancer model was explored [**7**]. 5.34µg/mL stigmasterol was required to inhibit the activity of TNF-α [**8**]. TNF-α is a major immune-modulatory and proinflammatory cytokine that is synthesized as a membrane-anchored precursor. It supports the previous studies indicating the anti-inflammatory properties of stigmasterol, with low concentrations of tumor necrosis factor-alpha (TNF-α). Monocytes and macrophages are responsible for releasing TNF-α, involved in the host immune response [**9**]. Pro-inflammatory mediators such as TNF- α, IL-6, and nitric oxide are crucial for the uninterrupted functioning of immune system especially in the presence of any infection. On the contrary, it can be the reason for the development of tissue and organ injury when overproduced [**10**]. Increased concentration of TNF-α is responsible for the pathogenic process of infectious as well as autoimmune diseases [**11**]. TNF- α is the mediator of tumor-associated inflammation and tumorigenesis [**12**]. It was observed as an important mediator of initiation and advancement of colitis-associated colon carcinogenesis [**3**]. Colon inflammation was found to be responsible for the promotion colon TNF- α which is linked to increased chances of colon pathogenesis [**13**].

Various studies associate tumor necrosis factor alpha converting enzyme (TACE) with colon cancer [**9**]. TACE is the member of A disintegrin and a metalloproteinase-containing enzyme (ADME). It is responsible for producing the soluble form of TNF-α from its membrane-bound precursor i.e. pro TNF-α. An experimental study conducted on teleost proved TNFα-converting activity of TACE [**14**]. Various other studies proved the involvement of TACE in the suppression of colon cancer [**9**]. It is well known that high concentration of TNF-α is found in the cases of colon cancer. Inhibition of TNF-α production can suppress the signals that are responsible for releasing the mature TNF-α. Although a number of proteases have been shown to process pro-TNF-alpha, the most efficient are TNF-α converting enzyme (TACE). Since the enzyme plays an important role in converting TNF-α in soluble form, targeting the enzyme could be a potential therapeutic strategy counteracting the increase in TNF-α concentration,

Copyright © 2017 VBRI Press

which has been seen in many cases of colon cancer. The present study explored a novel therapeutic perspective of stigmasterol for colon cancer in which overexpression of TNF-α was suppressed by stigmasterol through inhibition of TACE. Molecular docking studies have been used to identify the binding modes. A Molecular Dynamics (MD) simulation was carried out for the docked complex to analyze its stability inside the bodily conditions by using Gromacs simulation suite.

Experimental

Materials

Protein and ligand preparation

The crystal structure of TACE [PDB: 1BKC] was retrieved from Protein Data Bank. Structure of ligand molecule stigmasterol (CID: 5280794) was obtained from NCBI – PubChem Compound Database.

Prediction of active site

Understanding the reaction of catalytic residues is necessary in order to understand the accurate function of an enzyme. Although active site information of TACE was partially reported [**15**] earlier, current study presents more details on co-crystallized structure with its inhibitor and further validation has been done by *in silico* analysis. A most probable active site with amino acid residue information was obtained from Q-site Finder web server.

Molecular docking

Molecular docking was performed with AutoDock4 [16]. AutoDock4 is a one of the widely-used docking tool having an efficiency of predicting rapid and exact bound confirmations of ligand and their potential targets. AutoDock docking models are often consistent with X-ray crystal structures [**17, 18**].

MD simulation in water

A Molecular Dynamics (MD) simulation was carried out for the docked complex to analyze its stability inside the bodily conditions by using Gromacs simulation suite [**19**].

Methods

AutoDock uses a grid-based method for quick evaluation of binding energy of test confirmations. AutoDock starts with preprocessing of ligand by removal of coordinates from the PDB file and water molecules. AutoDock allows the target to be embedded in the grid while the investigating atom is consecutively directed to each grid point. Interaction energy of investigating atom and target molecule is calculated and is secured in the grid, which is then referred during docking simulation. Lamarckian genetic algorithm is used for conformational searching [**16**]. AutoDock4 uses a semi empirical free energy force field for the prediction of binding free energies of ligand to macromolecular targets. Q-site Finder anticipates the information active site residues by using energy criteria and computes van der Waals interactions of methyl probe with the target molecule. The clusters of favorable energies are graded according to the total interaction energies and the cluster of maximum energy filled the first rank [**20**]. Polar hydrogen was added to the protein structure to provide accurate ionization. The ligand was attributed to gasteiger charges and rigid roots. 13 Rotatable bonds were adjusted. Key residues were embedded in energy-container grid of 60 Å × 60 Å × 60 Å (x, y, z). The best confirmation was obtained through default Lamarckian genetic algorithm as a search protocol. The output from AutoDock4 was analyzed with Viewerlite 5.0. Ligplot was generated using PDBsum determined the number and length of H- bonds between the ligand and the receptor. The application of MD simulations to study the protein dynamics is well described in many previous studies [**21, 22, 23, 24**]. Here, we performed MD simulations for 30ns and studied the stability of protein as well as the ligand as a complex.

Result and Discussion

Identification of Active Catalytic site in TACE

Q-site Finder generated ten clusters based on probe and protein interactions after submitting the pre-processed structure of TACE. Individual probe site associate preferentially to the best-suited binding sites on the protein surface and these are the positions where a tentative ligand could interact and optimize its van der Waals interaction energy [20]. The obtained results included nearly 27 residues. All these residues are covered in the grid while performing docking. The active cleft of TACE has a catalytic zinc residue at its center pentacoordinated by the three-imidazole N_2 atoms of His 405, His 409, His 415. Along with these residues Glu 406, Met 345, Pro 437, Trp 312, Asp 344 around the zinc atom constitute the active site. The docked complex was forming a hydrogen bond with Gly 346 and showing hydrophobic interactions with many of the active site residues. **Fig. 1**(**a**) depicts the stigmasterol and TACE complex after molecular docking while the **Fig. 1**(**b**) is the ligplot showing hydrogen and hydrophobic bonds between ligand and macromolecule TACE.

Molecular docking simulation

We have investigated molecular docking and MD simulations on stigmasterol and TACE complex for the stability in the bodily conditions. CPK model is shown in **Fig. 2**. Computed Root mean square deviation (RSMD) (**Fig. 3**) suggests that docked complex is stable. Root mean square deviation was calculated by least square fitting of CA from the initial structure and found to be stable around 2A

Copyright © 2017 VBRI Press

during simulations; suggesting that the docked complex is stable.

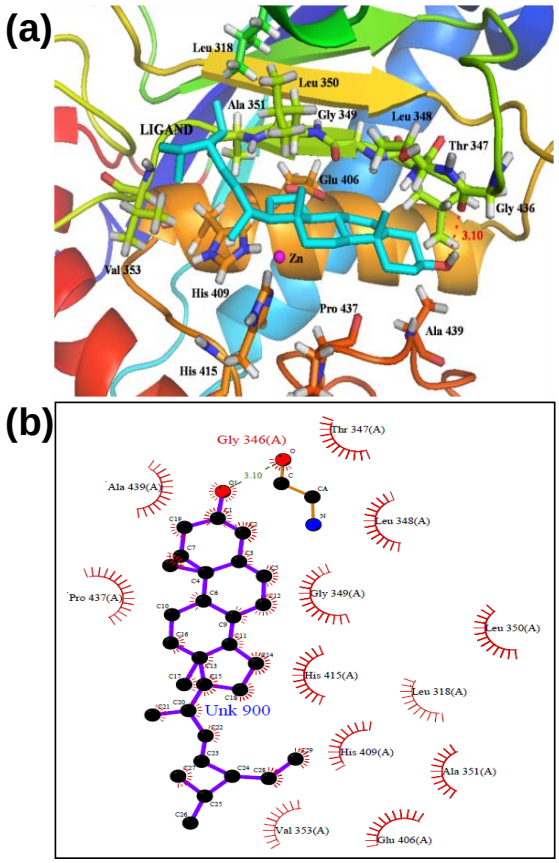


Fig. 1. (a)Stigmasterol and TACE Complex, (b) Ligplot showing hydrogen and hydrophobic bonds between ligand and macromolecule TACE.

Ligand and the active site residues in the receptor domain within 8A distance were considered for the interaction studies. A schematic of detailed ligand atom interactions with the protein residues before MD simulations is shown here (**Fig. 4**). The docked complex was forming a hydrogen bond with Gly 346 and within the hydrophobic interaction range with many of the active site residues. Hydrogen bond distance between ligand and receptor is shown in **Fig. 5** (**a**). Hydrogen bonding pattern between ligand and receptor is indicating the formation of strong hydrogen bond during simulations. Hydrogen bond existence map for various amino acids interacting with ligand is given in **Fig. 5** (**b**). Cyan implies the presence of a hydrogen bond and Red implies the absence of one. The *y*-coordinate shows the hydrogen-bond index.

In the present study, we examined the molecular mechanism of the stigmasterol on TACE that helps in understanding of the cancer cell characteristics. We observed that stigmasterol docked onto active site of TACE. The active site of TACE comprises zinc atom coordinated by a conserved zinc-binding domain (405-HexGHxxGxxH-415) [**15**]. It is also determined by the three-dimensional structure of TACE through Q-site Finder server. The ligand is interacting with key amino acids Leu 348 and Gly 349 present in the binding pocket of the enzyme [**25**].

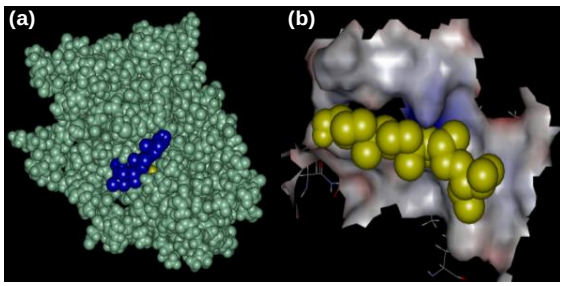


Fig. 2. complex1_tace rendered in cpk.

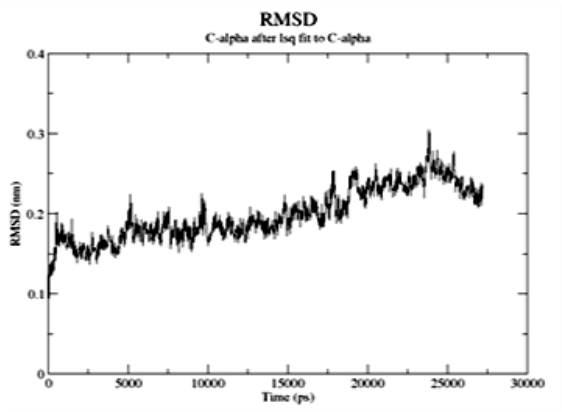


Fig. 3. Root mean square deviation.

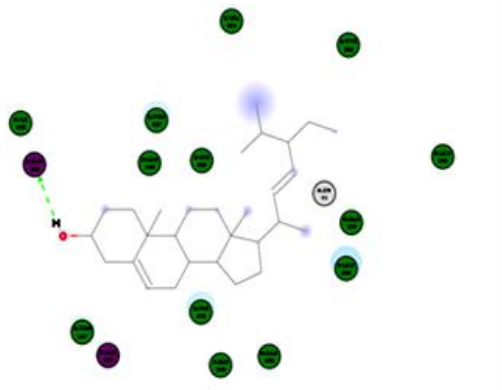


Fig. 4. A schematic representation of ligand interaction with the protein residues before performing MD simulations.

The binding pocket also co-occurred with the binding site of stigmasterol acquainted in the co-crystallized structure acquired from PDB. Molecular docking score of the 3D structure of stigmasterol with the above-described active site of TACE was -10.04 predicting good binding affinity of TACE with stigmasterol. Docked complex exhibited molecular interaction with glutamate and histidine residues that play a key function during the proteolytic reaction procedure. Short hairpin RNA silencing TACE (shTACE) prevented and efficiently treated acute and

Copyright © 2017 VBRI Press

chronic ulcerative colitis by decreasing TNF- α level [**26**]. The active site of TACE was blocked with highly selective new non-hydroxamate sulfonamide TACE inhibitors, considering as therapeutic targets for the treatment of TNF-dependent pathologies [**27**]. A quinazoline derivative therapeutically improved arthritis through suppressing production of TNF- α mediated by TACE [**28**]. Another study on TACE inhibition revealed TACE activation by non-receptor tyrosine kinase Src in mechanically stressed cardiomyocytes. This process could be harmful to specific blockade of TNF- α secretion, which in turn leads to congestive heart failure [**29**]. Considering above facts, the interaction of TACE and stigmasterol were anticipated to intervene with the interaction of substrate for the binding site of TACE, hence strengthening the idea of stigmasterol as TACE inhibitor.

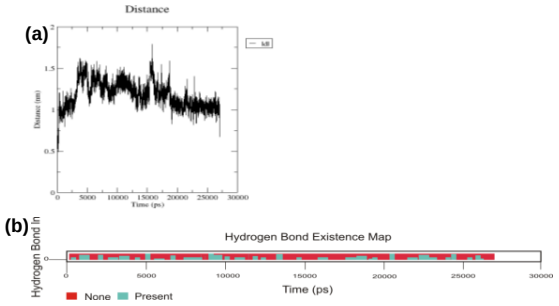


Fig. 5. (a): Hydrogen bond distance between ligand and receptor, (b) Hydrogen bond existence map.

Molecular dynamics simulation is a reliable and new way for the prevention of many diseases [**30**]. Molecular dynamics simulations are a widely used computer simulations technique to study the conformational dynamics of biological macromolecules such as proteins [**31**]. Molecular dynamics of the docked complex were examined to find out its stableness in bodily conditions. A simulation run time was 30ns, which was sufficient for the rearrangement of a ligand-bound protein molecule to acquire a stable binding manner. Root mean square deviation was calculated by least square fitting of CA from the initial structure and found to be stable around 2A during simulations; suggesting that the docked complex is stable. A structure showing the stationary phase was considered to study the molecular interaction pattern in the docked complex. H-bond, hydrophobic interaction, and van der Waals interaction were the factors that play an important role to make the stigmasterol and TACE a stable interaction. The dynamic stableness of stigmasterol during simulation runtime and its interaction with the important residues present at the binding site of the enzyme possibly consolidate the mode of action of stigmasterol on inhibition of TACE, which is responsible for the release of the soluble form of TNF-α.

Drug-likeliness

The successful evaluation has been done for drug-likeness based on Lipinski's "Rule of Five", while 46 computed physicochemical properties or molecular descriptors were used to predict ADMET (absorption, distribution, metabolism, elimination and toxicity) of the compound.

Conclusion

Molecular docking and simulation study revealed that presence of a ligand in the active site of TACE, interacting with the key residues consolidated the idea that stigmasterol is a potential inhibitor of TACE. Furthermore, this study is a stepping-stone in order understands the prevention mechanism of TNF-α, a high concentration of which is responsible for colon cancer. Hence, it can be used for further studies as a natural anti-cancer drug.

Acknowledgements

We would like to thank Vice Chancellor of Delhi Technological University, Head of Department of Department of Biotechnology and Institute of Biotechnology for providing the facility and support to conduct the research work.

Author's contributions

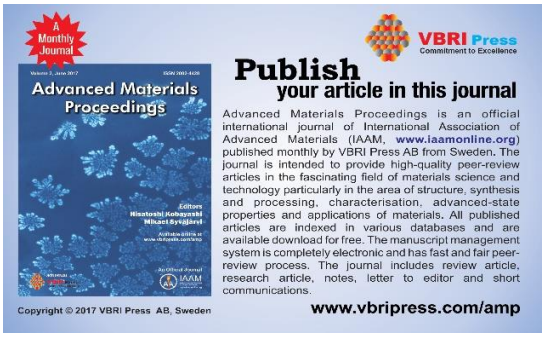
Conceived the plan: N.J., N.B., N.S.; Performed the experiments: N. J., N.S.; Data analysis: N.J., G.M., N.S., N.B; Wrote the paper: N.J., N.S., N.B. Authors have no competing financial interests.

References

1. American Cancer Society; Cancer Facts & Figures; Atlanta, Georgia. **2012**.
2. Kelloff, G. J.; Schilksy, R.L.; Alberts, D.S. et al. *Clin. Cancer Res*. **2004**, *10*, 3881. **DOI:** 10.1158/1078-0432.CCR-03-0783
3. Popivanova, B.K.; Kitamura, K.; Wu, Y.; Kondo,T.; Kagaya, T.; Kaneko, S.; Oshima, M.; Fujii, C.; Mukaida, N. J. *Clin. Invest*. **2008,** *118*, 560. **DOI:** 10.1172/JCI32453
4. Van Cutsem, E.; Verheul, H.M.; Flamen, P.; Rougier, P.; Beets-Tan, R.; Glynne-Jones, R.; Seufferlein, T.; *Cancers*, **2016,** *8*, 1. **DOI:** 10.3390/cancers8090081
5. Sievers, C.K.; Kratz, J.D.; Zurbriggen, L.D.; LoConte, N.K.; Lubnn, S.J.; Uboha, N.; Mulkerin, D.; Matkowskyj, K.A.; Deming, D.A.; *Clin. Colon. Rectal. Surg*., **2016,** *29*, 232. **DOI:** 10.1055/s-0036-1584292
6. Rizvi, M.Z.; Kukreja, A.K.; Bisht, N.S.*Physiology Mol. Biol. Plants*. **2015**, *21*, 617. **DOI:** 10.1007/s12298-015-0312-4
7. Ali, H.; Dixit, S.; Ali, D.; Alqahtani, S.M.; Alkahtani, S.;Alarifi, S.*Drug Des., Dev. Ther*.**2015**, 2793. **DOI:**10.2147/DDDT.S83514
8. Chao, W.W. and Lin, B.F. *Chin. Med*. **2010**, *5*, 17. **DOI:** 10.1186/1749-8546-5-17
9. Ramirez, J.A.; Bruttomesso, A. C.; Michelini, F. M.; Acebedo, S. L.; Alche, L. E.; Galagovsky, L. R. *Bioorg. Med. Chem.***2007**, *15*, 7538. **DOI:** 10.1016/j.bmc.2007.09.012
10. Szollosi, D. E.; Ghoneim, O.A.; Manzoor, M. K.; Dhuguru, J.; Edafiogho, I. O. *Inflammation*.**2016,** 1. **DOI:** 10.1007/s10753-016-0443-y

Copyright © 2017 VBRI Press

11. Magna, A. M.; Takiya, C. M.; Aruda, L. B.; Pascarelli, B.; Gomes, R. N.; Faria, H. C.; Lima, L. M.; Barreiro, E. J. *Int. Immonopharmacol*. **2005**, *5*, 485. **DOI:** 10.1016/j.intimp.2004.10.017
12. Grivennikov, S.I. and Karin, M. *Ann. Rheum. Dis*. **2011,** *70*, i104. **DOI:** 10.1136/ard.2010.140145
13. El-Bahrawy, A.; Tarhuni, A.; Kim, H.; Subramaniam, V.; Benslimane, I.; Abd Elmajeed, Z.Y.; Okpechi, S.; Ghonim, M.; Hemeida, R.; Abo-Yousef, A.; El-Sherbiny, G.; Abdel-Raheem, I.; Kim, J.; Naura, A.S.; Boulares, H. *Biosci. Rep*.**2016.** **DOI:** 10.1042/BSR20160195
14. Lu, X.J; Chen, Q.; Yang, G.J.; Chen, J. *Mol. Immunol*.**2015,** *63*, 497. **DOI:** 10.1016/j.molimm.2014.10.010
15. Maskos, K.; Femandez-catalan, C.; Huber, R.; Bouenkov, G.P.; Bartunik, H.; Ellestad, G.A.; Reddy, P.; Wolfson, M.F.; Rouch, C.T.; Castner, B.J.; Davis, R.; Clarke, H.R.; Petersen, M.; Fitzner, J.N.; Cerretti, D.P.; March, C.J.; Paxton, R.J.; Black, R.A; Bode, W. *Biochemistry*. **1998**, *95*, 3408. **DOI:** 10.1073/pnas.95.7.3408
16. Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. *J. Comput. Chem.* **1998**, *19*, 1639. **DOI:**10.1002/(SICI)1096-987X(19981115)19:14<1639:AID-JCC10>3.0.CO;2-B
17. Dym, O.; Xenarios, I.; Ke H.; Colicelli J. Mol. Pharmacol. **2002,***61*, 20. **DOI:**10.1124/mol.61.1.20
18. Rao, M.S. and Olson, A. J., *Proteins: Struct., Funct., Bioinf.* **1999**, *34*, 173. **DOI:** 10.1002/(SICI)1097-0134(19990201)34:2<173::AID-PROT3>3.0.CO;2-F
19. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J. *J. Comput. Chem.* **2005**, *26*, 1701. **DOI:** 10.1002/jcc.20291
20. Laurie, A.T.R. and Jackson, R.M. *Bioinformatics.***2005**, *21*, 1908. **DOI:** 10.1093/bioinformatics/bti315
21. Abe, C.; Benz, M.; Gajula, MNVP.; Vogel, K.P.; Gastel, M. V.; Illenberger, S.; Ziegler W.H.; Steinhoff, H.J.; *Biophys. J.***2011**, *101*, 1772. **DOI:** 10.1016/j.bpj.2011.08.048
22. Kumar, A.; Kumar, S.; Kumar, U.; Suravajhala, P., Gajula, M.N.V.P. *Comput. Bio.Chem.***2016**, *64*, 217.
23. Gajula, M.N.V.P.; Milikisyants, S.; Steinhoff, H.J.; Huber, M. *Appl. Magn. Reson*. **2007**, *31*, 99.
24. Gajula, M.N.V.P.; Soni, G.; Babu, G.; Bharadwaj, N. *J. Appl. Bioinform. Comput. Biol*. **2013**, *2*, 61. **DOI:** 10.4172/2329-9533.1000103
25. Elumalai, M.; Muthaiah, R.; Alf, M.A; *Acta. Pol. Pharm.***2012**, 69, 237. **DOI:** http://www.biomedcentral.com/1471-2164/14/S2/S4
26. Song, Y.; Kim, Y.R.; Kim, S.M.; UI Ain, Q.; Jang, K.; Yang, C.S.; Kim, Y.H. *J Controlled Release* **2016**, 239, 23. **DOI:** 10.1016/j.jconrel.2016.08.017
27. Sarkate, A.P.; Murumkar, P.R.; Lokwani, D.K.; Kandhare, A.D.; Bodhankar, S.L.; Shinde, D.B.; Bothara, K.G. *Environ. Res.***2015,** 26, 905. **DOI:** 10.1080/1062936X.2015.1095240.
28. Pu, Y.; Cao, D.; Xie, C.; Pei, H.; Li, D.; Tang, M.; Chen, L. *Biochem. Biophys. Res. Commun*.**2015**, 462, 288. **DOI:** 10.1016/j.bbrc.2015.04.111
29. Niu, A.; Wang, B.; Li, Y.P. *J. Cell. Biochem.***2015,** 11, 559. **DOI:** 10.1002/jcb.25006.
30. Gajula M.P., Vogel K.P., Rai A., Dietrich F. and Steinhoff H.J. (**2013**) BMC Genomics, 14(2), S4, 1-11 **DOI:** 10.1186/1471-2164-14-S2-S4
31. IV Borovykh, S Ceola, P Gajula, P. Gast, H. J. Steinhoff, Huber, M., *Journal of magnetic resonance*, **2006**, *180*, 178. **DOI:** 10.1016/j.jmr.2006.02.008
32. MNVP Gajula, A Kumar, Ijaq, J., *Protocol for Molecular Dynamics Simulations of Proteins*, **2016**, 6. **DOI:**10.21769/BioProtoc.2051
33. Ramprasad, E., *In silico analysis of chalcone synthase 1 protein sequences from different plant species. International Journal of Science, Environment and Technology*, **2016**, *5*, 1968.



J Ind Eng Int
DOI 10.1007/s40092-017-0212-z

ORIGINAL RESEARCH

Availability analysis of mechanical systems with condition-based maintenance using semi-Markov and evaluation of optimal condition monitoring interval

Girish Kumar[1] · Vipul Jain[2] · O. P. Gandhi[3]

Received: 17 April 2015 / Accepted: 8 June 2017
© The Author(s) 2017. This article is an open access publication

Abstract Maintenance helps to extend equipment life by improving its condition and avoiding catastrophic failures. Appropriate model or mechanism is, thus, needed to quantify system availability vis-a-vis a given maintenance strategy, which will assist in decision-making for optimal utilization of maintenance resources. This paper deals with semi-Markov process (SMP) modeling for steady state availability analysis of mechanical systems that follow condition-based maintenance (CBM) and evaluation of optimal condition monitoring interval. The developed SMP model is solved using two-stage analytical approach for steady-state availability analysis of the system. Also, CBM interval is decided for maximizing system availability using Genetic Algorithm approach. The main contribution of the paper is in the form of a predictive tool for system availability that will help in deciding the optimum CBM policy. The proposed methodology is demonstrated for a centrifugal pump.

Keywords Condition based maintenance · Availability · Semi-Markov process · Degraded states · Mechanical repairable systems

List of symbols

A_s	System availability
A_{pump}	Pump availability
D_i	ith degraded state, $i \in \{1, 2, 3, \ldots, n\}$
D_{x_i}	ith degraded state for CM metric x_i
F_r	Random failure state
$F_{ij}(t)$	CDF of the time spent in state i before moving to state j, $(i, j \in \Omega)$
$\bar{F}_{ij}(t)$	Complement of the $F_{ij}(t)$
F_i	CDF for CM interval for state i, where $i \in \{D_{x_1}, D_{x_2}, \ldots, D_{x_n}\}$
$\bar{F}_i$	Complement of the CDF for CM interval for state i, where $i \in \{D_{x_1}, D_{x_2}, \ldots, D_{x_n}\}$
im_p^r	Imperfect repair state where repair carried out to pth degraded state leads to rth better degraded state, $p \in \{3, \ldots, n\}$, $r = 1, 2, 3, \ldots,$ p-2
$K(t)$	SMP Kernel matrix
$k_{ij}(t)$	Element of the kernel matrix, $K(t)$ in ith row and jth column, $i, j \in \Omega$
l_i	Lower bound of CM metric for ith degraded state
$\mathrm{ma_k}$	Major repair state, where repair carried out at kth degraded state and leads to first degraded state, D_1, $k \in \{2, 3, \ldots, n\}$
mi_j	Minimal repair state corresponding at jth degraded state and maintenance leads to current degraded state, where $j \in \{1, 2, 3, \ldots, n\}$
n	Number of degradation states for the system
N	Total number states in the SMP model of the system
p_{ij}	(i, j)th element of one step transition probability matrix, Z

✉ Girish Kumar
girish.kumar154@gmail.com

Vipul Jain
vipul.jain@vuw.ac.nz

O. P. Gandhi
opgandhi_iitd@yahoo.com

[1] Department of Mechanical Engineering, Delhi Technological University, Bawana Road, Delhi 110042, India

[2] School of Management, Victoria Business School, Victoria University of Wellington, Wellington, New Zealand

[3] ITMMEC, Indian Institute of Technology Delhi, Hauz Khas, Delhi 110016, India

Published online: 19 June 2017

p_j	Probability of transitioning to state j from CM state and j is set of states having corresponding degraded state or possible maintenance state after CM
P_j	Steady state probability of the system being in working state j, $j \in$ W
P_i	Steady state probability of system being in state i, $i \in \Omega$
S	System
T_i	CM interval for degraded state D_i, $i \in \{1, 2, 3, \ldots, n\}$
t	Time
u(t-L)	Unit step function
u_i	Upper bound of CM metric for ith degraded state
W	Set of working states
Z	One step transitional probability matrix
τ_i	Mean sojourn time of the ith state, $i \in \Omega$
Ω	State space in the SMP model of system
β	Shape parameter of Weibull distribution
β_{ij}	Shape parameter of Weibull distribution for the time duration in ith state and related to state i to state j transition
θ_{ij}	Weibull distribution scale parameter for the time duration in ith state and related to state i to state j transition
θ	Weibull distribution scale parameter
μ_{ij}	Location parameter, mean, of log-normal distribution for the time duration in ith state and related to state i to state j transition
σ_{ij}	Scale parameter, standard deviation, of lognormal distribution for the time duration in ith state and related to state i to state j transition
λ	Parameter of the Exponential distribution
λ_i	Exponential distribution parameter of the CM time for state i, where $i \in \{C_1, C_2, \ldots, C_n\}$
H	Steady state probability matrix of the EMC
h_i	Steady state probability of ith state of the EMC, $i \in \Omega$

Introduction

Mechanical systems are employed in numerous applications, such as power plant, aviation and manufacturing. These demand higher availability for economic considerations. The analysis of availability guides practicing engineers in selecting an appropriate maintenance strategy and in improving performance of the system (Kumar et al. 2009).

In the past, researchers have focused towards development of time-based maintenance models (Crocker and Kumar 2000). Preventive maintenance (PM) is the scheduled maintenance that includes actions such as; adjustments, inspection, lubrication and replacement of components (Sim and Endrenyi 1988). Ahmad et al. (2011) proposed a maintenance decision model for PM application. Charles-Owaba et al. (2008) suggested scheduling of the PM based on opportunistic cost. Though PM does help in avoiding the equipment breakdowns, but it no way fulfills maintenance objectives by not utilizing optimum use of maintenance resources. The replacement interval of components and lube oils under PM is conservative, and this result in their replacement, with sufficient residual life still left. To take care of this concern, CBM policy came up. Under this; a planned maintenance action, is undertaken, i.e., not timely based but condition based. In CBM, condition-monitoring (CM) techniques such as vibration, acoustic emission, thermography and oil and wear debris analysis are used to assess health of the system (Wang 2008).

Most of the mechanical systems experience gradual degradation. It is appropriate to consider multiple states of the considered mechanical system or component, instead of the binary states for availability analysis (Majid and Nasir 2011). Van et al. (2012) studied maintenance policy with stochastic degradation of equipment. In degradation models; the states are decided, in general, based on the real-time health of the system. Wang et al. (2010) proposed a multi-degrading states model and considered the periodic inspection at each degraded state with an appropriate maintenance action.

Condition-based maintenance modeling for mechanical systems has been reviewed by various researchers (Jardine et al. 2006; Pandian and Ali 2009) with noticeable contributions from Chen and Trivedi (2005); Carr and Wang (2011). However, the transition distributions for degradation, repair and inspection are considered exponential, which is not always applicable. Amari and McLaughlin (2004) and El-Damcese and Temraz (2011) demonstrated the analysis of a multi-state system using Markov reward model and Markov chain, respectively. Ahmad et al. (2017) suggested a simulation-based optimization approach for free distributed repairable multi-state availability-redundancy allocation problems. Semi-Markov process approach is, however, capable to handle non-exponential distributions. But these have been attempted for limited applications such as software (Xie et al. 2005), NC machine tool (Qiang et al. 2010) and opportunistic maintenance (Kumar et al. 2014).

Although extensive work has been carried out under CBM, yet little attention has been paid to availability modeling of systems with CBM maintenance (Khanduja et al. 2011). Moreover, the frequency of CM interval affects the system availability and the literature shows that there is limited work on optimization of CM interval. In this paper, SMP modeling of mechanical systems, with multi-state degradation and using CBM, is proposed for the steady-state availability analysis.

The paper is organized, with system modeling described in the next Section. In Sect. 3, steady-state availability analysis is presented, while in Sect. 4, optimization of CM interval using GA is suggested and its methodology is described in Sect. 5. In Sect. 6, the proposed methodology is illustrated by an example and the last section, i.e., Sect. 7, concludes the paper.

System modeling

Condition-monitoring (CM) of the system is getting easier day-by-day, with recent advancements in sensor and real-time data capturing technologies. This has indeed helped to implement CBM in the plants in an effective way. In this work, a methodology for system availability assessment is developed, incorporating the CBM. The degradation behavior of the mechanical systems, together with periodic CM and their maintenance causes the stay time of their states that follow non-exponential distribution (Fricks and Trivedi 1997; Dennis 2003). The system model, therefore, is developed using the SMP, rather than a Markov model.

Modeling degradation, random failures and condition monitoring

In this subsection, three modeling aspects; degradation, random failures and periodic condition monitoring are considered. A SMP model of a mechanical system is developed employing the three features and is shown in Fig. 1.

Multi-degraded states with Weibull distribution are appropriate for developing the system model (Jager and Bertsche 2004). Let D_i, be a consecutive degraded state with $i \in \{1, 2, 3, \ldots, n\}$. Every degraded state, D_i, is identified using an appropriate CM metric, with lower and upper bounds; l_i and u_i. It is assumed that there is gradual system degradation; with the degradation rate slow initially, which increases with time.

Application of the CBM for a system does rule out the occurrence of its random failure. But this has been considered in this work to take care of an unforeseen case, i.e., the system may fail randomly in spite of being under CM. The random failures are modeled with exponential distribution and may occur at any stage of degradation. A random failure in the model, Fig. 1, is shown by connecting the degraded state, D_i, to random failure state, F_r by a directed line or edge, with direction from node D_i to node F_r. It may be mentioned that the system is brought back to its first state, D_1, by an appropriate repair action and the lognormal distribution is considered for such repairs (Dennis 2003).

In principle, the CM interval between two consecutive CM checks should be longer in the initial degraded states, while it should be shorter in the subsequent and high degradation states. Therefore, CM intervals; T_1, T_2, T_3, …, T_n, for the respective degraded states are considered in the decreasing order. Let D_{xi}, be the ith degraded state for CM metric, x_i, with $l_i \leq x_i \leq u_i$. The system may iterate within the same degraded state, D_i, and within its metric range, by adopting or following the minimal repair. However, as the system ages, it moves to the next degraded state. The time

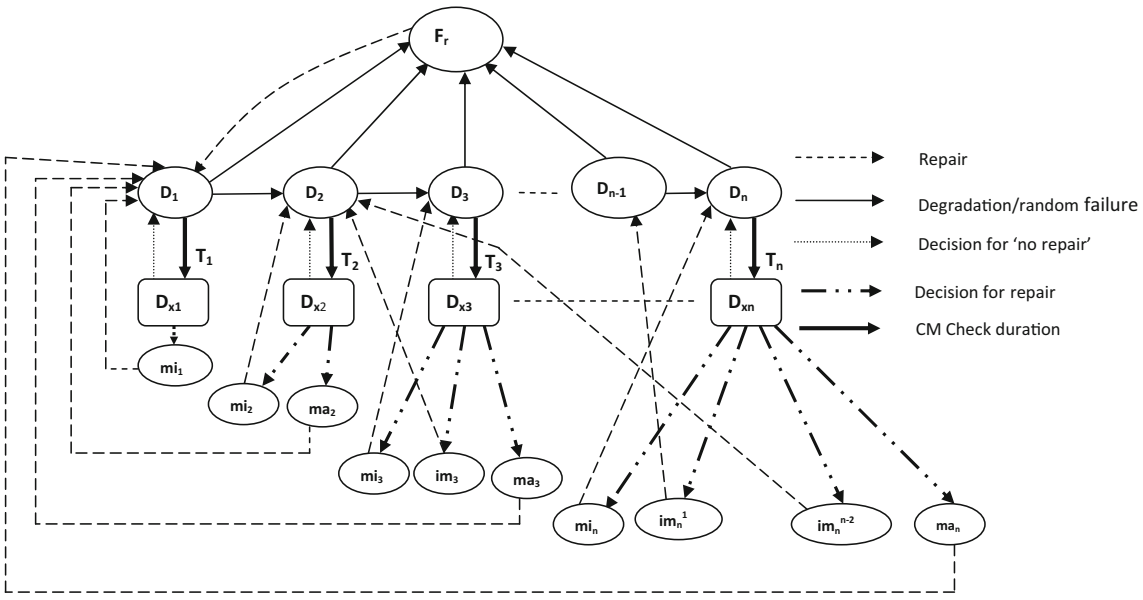


Fig. 1 System SMP model—CBM

to perform CM at a degraded state is assumed to follow an exponential distribution.

Modeling various repair actions

The decision whether to undertake maintenance or maintain the status quo, is taken based on the condition of the system. In case the system health is normal and does not show deterioration from the CM metric check, the decision is 'no repair' and the system is allowed to operate. Otherwise, an appropriate maintenance action such as: minimal, imperfect or major repair is carried out to bring back the system to its current, previous better or the initial state, respectively.

In the system SMP model, Fig. 1, minimal repair is considered for all the degraded states, i.e., mi_j,- while major repair is considered for all degraded states except the first degraded state, i.e., ma_k, $k \in \{2, 3, \ldots, n\}$ and imperfect repair is considered at all degraded states, except the first and second degraded states, i.e., im_p^r, $p \in \{3, \ldots, n\}$ and $r = 1, 2, 3, \ldots, (p\text{-}2)$. It is well known that a minimal repair takes, in general, the least time, while the imperfect repair takes more time than the minimal repair, but a lesser time than a major repair. Moreover, the time of repair will be higher, if the repair is carried out at a highly degraded state. The lognormal distribution, being more appropriate for repair of mechanical systems, is employed in the model (Dennis 2003).

The SMP system model, Fig. 1, is developed by incorporating the aspects discussed in this section. A solution procedure of the model for the system availability analysis is described in the next section.

Steady-state availability analysis

This section deals with steady state availability analysis based on the SMP system model in Sect. 2 and using analytical solution described in the following subsection.

Analytical solution of the SMP model

A two-stage method which is easy to implement, is selected for steady-state solution of the SMP model (Xie et al. 2005). In the two-stage method; Embedded Markov chain (EMC) of SMP that is a discrete time Markov chain (DTMC), is used to simplify the analysis of the SMP. Stage 1, deals with evaluation of the one-step transition probability matrix of Embedded Markov Chain of the SMP model, which is used to obtain steady-state probabilities of the EMC. In stage 2, the sojourn (stay) time of each state in a SMP model is evaluated and its steady-state probability is obtained using the values of the steady-state probabilities of the EMC (obtained from stage 1) and its sojourn time value. The procedure for both the stages is detailed in the following subsections.

Stage 1: steady-state probabilities of EMC

Refer Fig. 1 for the system SMP model, which consists of the state space, Ω, represented as:

$$\Omega = \{D_1, \ldots, D_n;\quad F_r;\quad D_{x_1}, \ldots, D_{x_n};\quad mi_1, \ldots, mi_n;\quad im_3^1, \ldots, im_n^{n-2};\quad ma_2, \ldots, ma_n\}$$

The Semi-Markov Process model is expressed in terms of kernel matrix, $K(t)$, which is given as:

$$K(t) = \begin{array}{c} \begin{array}{cccccccccccccccc} D_1 & \ldots & D_n & F_r & D_{x_1} & \ldots & D_{x_n} & mi_1 & \ldots & mi_n & im_3^1 & \ldots & im_n^{n-2} & ma_2 & \ldots & ma_n \end{array} \\ \begin{bmatrix} k_{D_1D_1} & \cdots & k_{D_1D_n} & k_{D_1F_r} & k_{D_1D_{x_1}} & \cdots & k_{D_1D_{x_n}} & k_{D_1mi_1} & \cdots & k_{D_1mi_n} & k_{D_1im_3^1} & \cdots & k_{D_1im_n^{n-2}} & k_{D_1ma_2} & \cdots & k_{D_1ma_n} \\ \vdots & \vdots\vdots & \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ k_{D_nD_1} & \cdots & k_{D_nDn} & k_{D_nF_r} & k_{D_nD_{x_1}} & \cdots & k_{D_nD_{x_n}} & k_{D_nmi_1} & \cdots & k_{D_nmi_n} & k_{D_nim_3^1} & \cdots & k_{D_nim_n^{n-2}} & k_{D_nma_2} & \cdots & k_{D_nma_n} \\ k_{F_rD_1} & \cdots & k_{F_rD_n} & k_{F_rF_r} & k_{F_rD_{x_1}} & \cdots & k_{F_rD_{x_n}} & k_{F_rmi_1} & \cdots & k_{F_rmi_n} & k_{F_rim_3^1} & \cdots & k_{F_rim_n^{n-2}} & k_{F_rma_2} & \cdots & k_{F_rma_n} \\ k_{D_{x_1}D_1} & \cdots & k_{D_{x_1}D_n} & k_{D_{x_1}F_r} & k_{D_{x_1}D_{x_1}} & \cdots & k_{D_{x_1}D_{x_n}} & k_{D_{x_1}mi_1} & \cdots & k_{D_{x_1}mi_n} & k_{D_{x_1}im_3^1} & \cdots & k_{D_{x_1}im_n^{n-2}} & k_{D_{x_1}ma_2} & \cdots & k_{D_{x_1}ma_n} \\ \vdots & \vdots\vdots & \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ k_{D_{x_n}D_1} & \cdots & k_{D_{x_n}D_n} & k_{D_{x_n}F_r} & k_{D_{x_n}D_{x_1}} & \cdots & k_{D_{x_n}D_{x_n}} & k_{D_{x_n}mi_1} & \cdots & k_{D_{x_n}mi_n} & k_{D_{x_n}im_3^1} & \cdots & k_{D_{x_n}im_n^{n-2}} & k_{D_{x_n}ma_2} & \cdots & k_{D_{x_n}ma_n} \\ k_{mi_1D_1} & \cdots & k_{mi_1D_n} & k_{mi_1F_r} & k_{mi_1Dx_1} & \cdots & k_{mi_1D_{x_n}} & k_{mi_1mi_1} & \cdots & k_{mi_1mi_n} & k_{mi_1im_3^1} & \cdots & k_{mi_1im_n^{n-2}} & k_{mi_1ma_2} & \cdots & k_{mi_1ml_n} \\ \vdots & \vdots\vdots & \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ k_{mi_nD_1} & \cdots & k_{mi_nD_n} & k_{mi_nF_r} & k_{mi_nDx_1} & \cdots & k_{mi_nD_{x_n}} & k_{mi_nmi_1} & \cdots & k_{mi_nmi_n} & k_{mi_nim_3^1} & \cdots & k_{mi_nim_n^{n-2}} & k_{mi_nma_2} & \cdots & k_{mi_nma_n} \\ k_{im_3^1D_1} & \cdots & k_{im_3^1D_n} & k_{im_3^1F_r} & k_{im_3^1D_{x_1}} & \cdots & k_{im_3^1D_{x_n}} & k_{im_3^1mi_1} & \cdots & k_{im_3^1mi_n} & k_{im_3^1im_3^1} & \cdots & k_{im_3^1im_n^{n-2}} & k_{im_3^1ma_2} & \cdots & k_{im_3^1ma_n} \\ \vdots & \vdots\vdots & \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ k_{im_n^{n-2}D_1} & \cdots & k_{im_n^{n-2}D_n} & k_{im_n^{n-2}F_r} & k_{im_n^{n-2}D_{x_1}} & \cdots & k_{im_n^{n-2}D_{x_n}} & k_{im_n^{n-2}mi_1} & \cdots & k_{im_n^{n-2}mi_n} & k_{im_n^{n-2}im_3^1} & \cdots & k_{im_n^{n-2}im_n^{n-2}} & k_{im_n^{n-2}ma_2} & \cdots & k_{im_n^{n-2}ma_n} \\ k_{ma_2D_1} & \cdots & k_{ma_2D_n} & k_{ma_2F_r} & k_{ma_2D_{x_1}} & \cdots & k_{ma_2D_{x_n}} & k_{ma_2mi_1} & \cdots & k_{ma_2mi_n} & k_{ma_2im_3^1} & \cdots & k_{ma_2im_n^{n-2}} & k_{ma_2ma_2} & \cdots & k_{ma_2ma_n} \\ \vdots & \vdots\vdots & \vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots & \vdots & \vdots\vdots & \vdots \\ k_{ma_nD_1} & \cdots & k_{ma_nD_n} & k_{ma_nF_r} & k_{ma_nD_{x_1}} & \cdots & k_{ma_nD_{x_n}} & k_{ma_nmi_1} & \cdots & k_{ma_nmi_n} & k_{ma_nim_3^1} & \cdots & k_{ma_nim_n^{n-2}} & k_{ma_nma_2} & \cdots & k_{ma_nma_n} \end{bmatrix} \begin{array}{c} D_1 \\ \vdots \\ D_n \\ F_r \\ D_{x_1} \\ \vdots \\ D_{x_n} \\ mi_1 \\ \vdots \\ mi_n \\ im_3^1 \\ \vdots \\ im_n^{n-2} \\ ma_1 \\ \vdots \\ ma_n \end{array} \end{array} \tag{1}$$

The elements, $k_{ij}(t)$ of the matrix, $K(t)$, as given by Kulkrani (2009) are redefined as:

$$k_{ij}(t) = \begin{cases} 0 & \text{: when transition is not possible from the ith state within the time } t \\ F_{ij}(t) & \text{: when only one state j is reachable from the ith state (excluding the CM states) within the time } t \\ \int_0^t \overline{F}_{ik}\overline{F}_{im}\mathrm{d}F_{ij}(x) & \text{: when two or more states (for example; three states viz. } j, k \text{and } m) \text{ are reachable from the ith state (excluding the CM state) within the} \\ p_j F_i & \text{:when more than one state (i.e. degraded state and possible maintenance states) reachable from the CM state } i,\ \mathrm{i} \in \{D_{x1}, D_{x2}, \ldots, D_{xn}\} \text{ within time t} \end{cases} \quad (2)$$

The Kernel matrix, $K(t)$, Eq. (2), is used to evaluate one-step transition probability matrix, $Z = K(\infty)$, of the EMC of the SMP by considering $t \to \infty$. Let the elements of the matrix, Z, be p_{ij}, with $i, j = 1$–N. To evaluate these elements, there is an additional condition that for each row, the sum of elements of $K(t)$ becomes 1, as $t \to \infty$, i.e., $\sum_{j=1}^{N} p_{ij} = 1$ for all $i = 1$–N. Since these elements involve complicated integrals for the chosen CDF, these can be solved with the help of a software like Maple. The one-step transition probability matrix, Z, is given as (Kulkrani 2009):

$$Z = K(\infty) = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & \cdots & p_{1N} \\ p_{21} & p_{22} & p_{23} & \cdots & \cdots & p_{2N} \\ p_{31} & p_{32} & p_{33} & \cdots & \cdots & p_{3N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{N1} & p_{N2} & p_{N3} & \cdots & \cdots & p_{NN} \end{bmatrix} \quad (3)$$

The following system of linear equations is formulated to find the steady state probabilities of the EMC, $h_i, i \in \Omega$, (Kulkrani 2009).

$$H = HK(\infty), \quad \sum_{i=1}^{N} h_i = 1, i \in \Omega \quad (4)$$

where $H = [h_1, h_2, \ldots, h_N]$ is the steady state probability matrix of the EMC and $K(\infty)$ is the Kernel matrix when $t \to \infty$.

The system of linear Equations, i.e., Eq. (4), can be solved by software such as Matlab. The steady state probabilities of the EMC, h_i, $i \in \Omega$, are obtained, and these values are used for evaluating steady state probability, P_i, $i \in \Omega$ of the SMP of state i, in stage 2.

Stage 2: steady-state probabilities of states in SMP

In stage 2, the mean sojourn time, τ_i, i.e., the time the process spends at state i is evaluated to obtain the steady-state probability of state i, i.e.; P_i. The mean sojourn time, τ_i, is obtained using the following definition (Kulkrani 2009):

$$\tau_i = \begin{cases} 0 & \text{: when transition is not possible from the ith state within the time } t \\ \int_0^\infty \bar{F}_{ij}(t)\mathrm{d}t & \text{: when only one state } j \text{ is reachable from the ith state (excluding the CM states) within the time } t \\ \int_0^\infty \overline{F}_{ij}\bar{F}_{ik}\bar{F}_{im}(t)\mathrm{d}t & \text{: when two or more states (for example; three states viz. } j,\ k \text{ and } m) \text{ are reachable from the } i\text{th state (excluding the CM state) within the time } t \text{ with three respective elements of } k(t) \text{ in the } i\text{th row} \\ \int_0^\infty \overline{F}_{D_{x_i}}\mathrm{d}t & \text{: when more than one state (i.e. degraded state and possible maintenance states) are reachable from the CM state } i \\ & i \in \{D_{x1}, D_{x2}, \ldots, D_{xn}\} \text{ within time } t \end{cases} \quad (5)$$

Using the values of steady state probabilities of the EMC and the sojourn time values, the steady state probability of state i, P_i, for the SMP model is obtained as (Kulkrani 2009):

$$P_i = \frac{h_i \tau_i}{\sum_{j \in \Omega} h_j \tau_j}, \quad i \in \Omega, \tag{6}$$

where h_i and τ_i values are evaluated using Eqs. (4) and (5), respectively

If W is set of working/operating state, the steady state availability is given by:

$$A_s = \sum_{j \in W} P_j \tag{7}$$

where, P_j is obtained using the Eq. (6).

Optimization of CM interval using GA

A genetic algorithm (GA) solves complex optimization problems and it is based on the natural evolution process (Goldberg 2006). In this approach, population changes continuously through cross breeding, mutation and natural selection. An explicit mathematical formulation is not required in implementation and, the parameter values, such as; population size, crossover rate and mutation rate are appropriately chosen to maintain the desired accuracy in the solution. The main steps of GA approach are (Kuo et al. 2001):

- Chromosome representation
- Generation of initial population
- Evaluation of fitness function
- Selection process
- Genetic operations (crossover, mutation)
- Selection of best chromosomes according to the fitness values
- Termination criteria

In this paper, the GA approach is applied to decide the optimal CM interval for maximizing the system availability. The availability function as per the procedure detailed in Sect. 3 is coded in Matlab and the developed code is used in the GA tool box for optimization.

A methodology based on the deliberations in Sects. 2, 3 and 4 is suggested for steady state system availability in the following section.

Step wise methodology: availability analysis

Steps for the steady-state availability analysis of the SMP model are presented below, which are based on the discussion in Sects. 2, 3 and 4.

Step I: system structure

Identify the components/subsystems in the selected system.

Step II: SMP modeling of the system

Derive feasible states for the system and evolve a SMP model (Refer Sect. 2).

Step III: component/subsystem failure and repair distribution

Identify the distribution and its parameter (s) for degradation/failure time, repair time, CM time interval and CM activity time at the component/subsystem level (Refer Sect. 2).

Step IV: two-stage steady state system availability analysis

(a) Kernel matrix, K(t)

Derive the Kernel matrix, $K(t)$, for the system model using the definition of its elements, Eq. (2).

(b) One step transition probability matrix, Z, of the EMC

Generate one-step transition probability matrix, Z, of the EMC of the SMP model, using the step IV (a), with $t \to \infty$.

(c) Steady state probability of the states of EMC

Steady state probability of all states of EMC of the system SMP model, using Eq. (4).

(d) Mean sojourn time of the states

Evaluate mean sojourn time of all states of the SMP model, using Eq. (5).

(e) Steady state probability of states

*St*eady state probability of state i, P_i, of the SMP model is obtained using the Eq. (6).

(f) Availability evaluation

Using step IV (e), evaluate the system steady state availability by adding the steady-state probabilities of the working states in the system, i.e., Eq. (7).

Step V: optimization of CM interval

Evaluate optimum CM interval to maximize the system availability by Genetic Algorithm approach (Refer Sect. 4).

The methodology suggested above is illustrated in the next section.

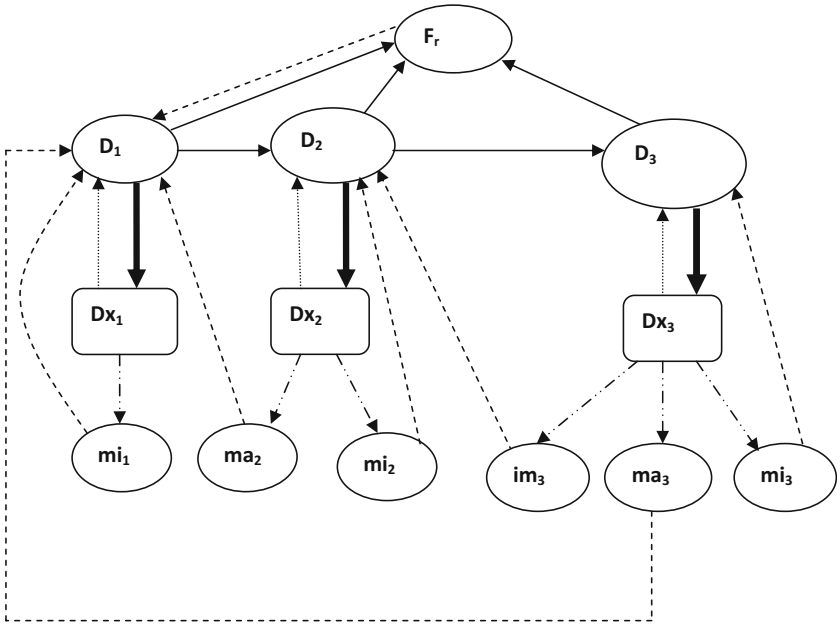


Fig. 2 SMP model; CBM—Centrifugal pump

Table 1 Distribution and parameter values—Centrifugal pump

State transition	Distribution	Parameter value	State transition	Distribution	Parameter value
$D_1 \rightarrow D_2$	Weibull	$\beta_{D_1D_2} = 1.1$, $\theta_{D_1D_2} = 15000$	$D_{x2} \rightarrow mi_2$	Exponential	$p_{mi_2} = 0.2,\ \lambda_{D_{x_2}} = \frac{1}{2}$
$D_1 \rightarrow F_r$	Exponential	$\lambda_{D_1F_r} = \frac{1}{30000}$	$D_{x2} \rightarrow ma_2$	Exponential	$p_{ma_2} = 1 - \{p_{D_2} + p_{mi_2}\} = 0.1$, $\lambda_{D_{x_2}} = \frac{1}{2}$
$D_1 \rightarrow X_1$	Deterministic	$T_1 = 240$	$D_{x3} \rightarrow D_3$	Exponential	$p_{D_3} = 0.4,\ \lambda_{D_{x_3}} = \frac{1}{3}$
$D_2 \rightarrow D_3$	Weibull	$\beta_{D_2D_3} = 1.5$, $\theta_{D_2D_3} = 11000$	$D_{x3} \rightarrow mi_3$	Exponential	$p_{mi_3} = 0.1,\ \lambda_{D_{x_3}} = \frac{1}{3}$
$D_2 \rightarrow F_r$	Exponential	$\lambda_{D_2F_r} = \frac{1}{25000}$	$D_{x3} \rightarrow im_3^1$	Exponential	$p_{im_3^1} = 0.2,\ \lambda_{D_{x_3}} = \frac{1}{3}$
$D_2 \rightarrow D_{X2}$	Deterministic	$T_2 = 168$	$D_{x3} \rightarrow ma_3$	Exponential	$p_{ma_3} = 1 - \{p_{D_3} + p_{mi_3} + p_{im_3^1}\} = 0.3,\ \lambda_{D_{x_3}} = \frac{1}{3}$
$D_3 \rightarrow F_r$	Exponential	$\lambda_{D_3F_r} = \frac{1}{20000}$	$mi_1 \rightarrow D_1$	Log-normal	$\mu_{mi_1D_1} = 1.4,\ \sigma_{mi_1D_1} = 0.41$
$D_3 \rightarrow D_{X3}$	Deterministic	$T_3 = 96$	$mi_2 \rightarrow D_2$	Log-normal	$\mu_{mi_2D_2} = 1.79,\ \sigma_{mi_2D_2} = 0.53$
$F_r \rightarrow D_1$	Log-normal	$\mu_{F_rD_1} = 1.95$, $\sigma_{F_rD_1} = 0.56$	$mi_3 \rightarrow D_3$	Log-normal	$\mu_{mi_3D_3} = 2.29,\ \sigma_{mi_3D_3} = 0.62$
$D_{x1} \rightarrow D_1$	Exponential	$p_{D_1} = 0.9,\ \lambda_{D_{x_1}} = 1$	$im_3^1 \rightarrow D_2$	Log-normal	$\mu_{im_3^1D_2} = 2.30,\ \sigma_{im_3^1D_2} = 0.64$
$D_{x1} \rightarrow mi_1$	Exponential	$p_{mi_1} = 1 - p_{D_1} = 0.1$, $\lambda_{D_{x_1}} = 1$	$ma_2 \rightarrow D_1$	Log-normal	$\mu_{ma_2D_1} = 2.08,\ \sigma_{ma_2D_1} = 0.59$
$D_{x2} \rightarrow D_2$	Exponential	$p_{D_2} = 0.7,\ \lambda_{D_{x_2}} = \frac{1}{2}$	$ma_3 \rightarrow D_1$	Log-normal	$\mu_{ma_3D_1} = 2.48,\ \sigma_{ma_3D_1} = 0.69$

Example

A centrifugal pump is selected for steady state availability assessment using the proposed methodology involving CBM model. Condition monitoring is applied for degradation/failure detection of the pump so that a maintenance action is taken well in time to avoid the failure of the pump and increase its availability. The suggested approach is demonstrated by following its step-by-step methodology given in Sect. 5.

Step I: system description

Pump system is considered. The centrifugal pump consists of six components, i.e., impeller, shaft, bearing, motor, bush, housing. To avoid the state space explosion problem, the states are defined at system level, i.e., at the pump level. The Semi-Markov model is developed for the system, i.e., the pump in the following step.

Step II: SMP model of the pump

Round the clock or continued operation of the pump can lead to its gradual deterioration. For the availability modeling, three degraded states are considered. Let these are: D_1, D_2 and D_3. A random failure state is designated as F_r. It is assumed that CM is performed at scheduled periodic intervals of T_1 for state D_1, T_2 for state D_2 and T_3 for state D_3. Let the corresponding CM metric for these degraded states are D_{x_1}, D_{x_2}, and D_{x_3}.. The maintenance activities are performed depending upon the condition of the pump. The minor repair is considered at each stage of degradation, with three minor repair states being mi_1,mi_2, and mi_3.. Major repair states, ma_2 and ma_3, are considered at the second and third stage of degradation. An imperfect repair state im_3^1 is envisaged at stage 3 of degradation. This is as per the discussion in Sect. 2.2.

Based on the above and referring Sect. 5.3, the SMP model is developed and is shown in Fig. 2. The state space for the model is given as:

$$\Omega = \{D_1, D_2, D_3; F_r; D_{x_1}, D_{x_2}, D_{x_3}; mi_1, mi_2, mi_3; im_3^1; ma_2, ma_3\}$$

The steady state analytical solution of the SMP model requires the state transition time distributions and their parameters, which are identified in the next step.

Step III: state transition time distribution and parameters of the centrifugal pump

The state transition time distribution and parameters for the centrifugal pump are chosen based on literature and also the experience (Nel and Haarhoff 2011; Dennis 2003). The distribution and its parameter(s) for the time duration in each state are identified and are given in Table 1.

Step IV: two-stage method for steady state analysis of Centrifugal pump

There are two stages, i.e., stage 1 and stage 2. The stage 1 of the two-stage method is as below:

(a) Kernel matrix, $K(t)$

The Semi-Markov model is expressed by its Kernel matrix, $K(t)$ and their elements are calculated as per the method given in Sect. 3.1.1. The Kernel matrix $K(t)$ for the pump is obtained and is given as the matrix, Eq. (8), shows its zero and other leftover non-zero elements, i.e., $k_{ij}(t)$.

$$K(t) = \begin{array}{c} \begin{array}{cccccccccccccccc} D_1 & .. & D_n & F_r & D_{x_1} & .. & D_{x_n} & mi_1 & .. & mi_n & im_3^1 & .. & im_n^{n-2} & ma_2 & .. & ma_n \end{array} \\ \begin{bmatrix} k_{D_1D_1} & .. & k_{D_1D_n} & k_{D_1F_r} & k_{D_1D_{x_1}} & .. & k_{D_1D_{x_n}} & k_{D_1mi_1} & .. & k_{D_1mi_n} & k_{D_1im_3^1} & .. & k_{D_1im_n^{n-2}} & k_{D_1ma_2} & .. & k_{D_1ma_n} \\ \vdots & :: & \vdots & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots \\ k_{D_nD_1} & .. & k_{D_nD_n} & k_{D_nF_r} & k_{D_nD_{x_1}} & .. & k_{D_nD_{x_n}} & k_{D_nmi_1} & .. & k_{D_nmi_n} & k_{D_nim_3^1} & .. & k_{D_nim_n^{n-2}} & k_{D_nma_2} & .. & k_{D_nma_n} \\ k_{F_rD_1} & .. & k_{F_rD_n} & k_{F_rF_r} & k_{F_rD_{x_1}} & .. & k_{F_rD_{x_n}} & k_{F_rmi_1} & .. & k_{F_rmi_n} & k_{F_rim_3^1} & .. & k_{F_rim_n^{n-2}} & k_{F_rma_2} & .. & k_{F_rma_n} \\ k_{D_{x_1}D_1} & .. & k_{D_{x_1}D_n} & k_{Dx_1F_r} & k_{D_{x_1}D_{x_1}} & .. & k_{D_{x_1}D_{x_n}} & k_{D_{x_1}mi_1} & .. & k_{D_{x_1}mi_n} & k_{D_{x_1}im_3^1} & .. & k_{D_{x_1}im_n^{n-2}} & k_{D_{x_1}ma_2} & .. & k_{D_{x_1}ma_n} \\ \vdots & :: & \vdots & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots \\ k_{D_{x_n}D_1} & .. & k_{D_{x_n}D_n} & k_{D_{x_n}F_r} & k_{D_{x_n}D_{x_1}} & .. & k_{D_{x_n}D_{x_n}} & k_{D_{x_n}mi_1} & .. & k_{D_{x_n}mi_n} & k_{D_{x_n}im_3^1} & .. & k_{D_{x_n}im_n^{n-2}} & k_{D_{x_n}ma_2} & .. & k_{D_{x_n}ma_n} \\ k_{mi_1D_1} & .. & k_{mi_1D_n} & k_{mi_1F_r} & k_{mi_1Dx_1} & .. & k_{mi_1D_{x_n}} & k_{mi_1mi_1} & .. & k_{mi_1mi_n} & k_{mi_1im_3^1} & .. & k_{mi_1im_n^{n-2}} & k_{mi_1ma_2} & .. & k_{mi_1ml_n} \\ \vdots & :: & \vdots & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots \\ k_{mi_nD_1} & .. & k_{mi_nD_n} & k_{mi_nF_r} & k_{mi_nDx_1} & .. & k_{mi_nD_{x_n}} & k_{mi_nmi_1} & .. & k_{mi_nmi_n} & k_{mi_nim_3^1} & .. & k_{mi_nim_n^{n-2}} & k_{mi_nma_2} & .. & k_{mi_nma_n} \\ k_{im_3^1D_1} & .. & k_{im_3^1D_n} & k_{im_3^1F_r} & k_{im_3^1D_{x_1}} & .. & k_{im_3^1D_{x_n}} & k_{im_3^1mi_1} & .. & k_{im_3^1mi_n} & k_{im_3^1im_3^1} & .. & k_{im_3^1im_n^{n-2}} & k_{im_3^1ma_2} & .. & k_{im_3^1ma_n} \\ \vdots & :: & \vdots & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots \\ k_{im_n^{n-2}D_1} & .. & k_{im_n^{n-2}D_n} & k_{im_n^{n-2}F_r} & k_{im_n^{n-2}D_{x_1}} & .. & k_{im_n^{n-2}D_{x_n}} & k_{im_n^{n-2}mi_1} & .. & k_{im_n^{n-2}mi_n} & k_{im_n^{n-2}im_3^1} & .. & k_{im_n^{n-2}im_n^{n-2}} & k_{im_n^{n-2}ma_2} & .. & k_{im_n^{n-2}ma_n} \\ k_{ma_2D_1} & .. & k_{ma_2D_n} & k_{ma_2F_r} & k_{ma_2D_{x_1}} & .. & k_{ma_2}D_{x_n} & k_{ma_2mi_1} & .. & k_{ma_2mi_n} & k_{ma_2im_3^1} & .. & k_{ma_2im_n^{n-2}} & k_{ma_2ma_2} & .. & k_{ma_2ma_n} \\ \vdots & :: & \vdots & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots & \vdots & :: & \vdots \\ k_{ma_nD_1} & .. & k_{ma_nD_n} & k_{ma_nF_r} & k_{ma_nD_{x_1}} & .. & k_{ma_n}D_{x_n} & k_{ma_nmi_1} & .. & k_{ma_nmi_n} & k_{ma_nim_3^1} & .. & k_{ma_nim_n^{n-2}} & k_{ma_nma_2} & .. & k_{ma_nma_n} \end{bmatrix} \end{array} \begin{array}{c} D_1 \\ \vdots \\ D_n \\ F_r \\ D_{x_1} \\ \vdots \\ D_{x_n} \\ mi_1 \\ \vdots \\ mi_n \\ im_3^1 \\ \vdots \\ im_n^{n-2} \\ ma_1 \\ \vdots \\ ma_n \end{array} \tag{8}$$

Table 2 Steady state probability of the states of the EMC—CBM model

h_i	Probability values	h_i	Probability values
h_{D_1}	0.4579649417	h_{mi_1}	0.0449538387
h_{D_2}	0.0156699944	h_{mi_2}	0.0077679296
h_{D_3}	0.0000583264	h_{mi_3}	0.0000087070
h_{F_r}	0.0037367736	$h_{im_3^1}$	0.000174141
$h_{D_{x_1}}$	0.4495383868	h_{ma_2}	0.0046607578
$h_{D_{x_2}}$	0.0155358593	h_{ma_3}	0.0000290235
$h_{D_{x_3}}$	0.0000580470		

Table 3 Steady state probability—CBM model

P_i	Probability values	P_i	Probability values
P_{D_1}	0.9694212251	P_{mi_1}	0.0017644151
P_{D_2}	0.0233338358	P_{mi_2}	0.0004763431
P_{D_3}	0.0000497153	P_{mi_3}	0.0000009276
P_{F_r}	0.0002733783	$P_{im_3^1}$	0.0000018970
$P_{D_{x_1}}$	0.0040009413	P_{ma_2}	0.0003953164
$P_{D_{x_2}}$	0.0002765417	P_{ma_3}	0.0000039134
$P_{D_{x_3}}$	0.0000015499		

The non-zero elements of the matrix, $K(t)$, Eq. (8), are derived, using the distributions (Table 1) and Eq. (2), and these are listed in "Appendix 1".

(b) One step transition probability matrix, Z, of the EMC

The Kernel matrix $K(t)$, Eq. (8), evaluates one-step transition probability matrix, $Z = K(\infty)$, of the EMC of the SMP, with $t \to \infty$. Refer Sect. 3.1.1. The elements of matrix Z are evaluated with the condition that for each row, the summation of elements of $K(t)$ is 1.

Using expressions of elements of the Kernel matrix, $K(t)$, listed in "Appendix 1", the elements of the matrix, Z are derived and these are tabulated in "Appendix 2".

It may be mentioned that the Matlab software is employed to evaluate the expressions (Matlab Release 7.7 2010). The evaluated non-zero elements of the matrix, Z, are as listed in "Appendix 2".

(c) Steady state probability of the states of the EMC

The steady state probabilities of the EMC are evaluated using a set of Eq. (4) for the system data (Table 1).

For the model, the system of equations is:

$$\begin{bmatrix} h_{D_1} & h_{D_2} & h_{D_3} & h_{F_r} & h_{D_{x_1}} & h_{D_{x_2}} & h_{D_{x_3}} & h_{mi_1} & h_{mi_2} & h_{mi_3} & h_{im_3^1} & h_{ma_2} & h_{ma_3} \end{bmatrix} = \begin{bmatrix} h_{D_1} & h_{D_2} & h_{D_3} & h_{F_r} & h_{D_{x_1}} & h_{D_{x_2}} & h_{D_{x_3}} & h_{mi_1} & h_{mi_2} & h_{mi_3} & h_{im_3^1} & h_{ma_2} & h_{ma_3} \end{bmatrix} * Z \quad (9)$$

This set of equations is solved using Matlab (Matlab Release 7.7 2010) and the values of $h_i, i \in \Omega$ are tabulated and given in Table 2.

(d) Mean sojourn time of the states of the centrifugal pump

Sojourn time values are evaluated for the stage 2 and these values are used for evaluation of steady state probabilities of system states. Sojourn time, τ_i, expressions are derived using Eq. (5) and using the distribution parameter value given in Table 1. These expressions are listed in "Appendix 3". Using the set of expressions ("Appendix 3") and the distribution and its parameter values, given in Table 1, the sojourn time for all states, involving complex integrals, is evaluated (Matlab Release 7.7 2010) and the same are listed in "Appendix 3".

(e) Steady state probability of states of the centrifugal pump

The steady state probability of state i, for the pump system, is evaluated using Eq. (6), and the values of steady state probabilities of EMC, h_i, and the values sojourn times (τ_i) as calculated above in step IV (c) and IV (d). The evaluated value of the steady state probability of states is given in Table 3.

(f) System steady state availability for the centrifugal pump

System steady state availability assessment of the pump system is obtained using Eq. (7), i.e.,

$$A_{\text{pump}} = \sum_{j \in W} P_j$$

The set of working states of the system model (Fig. 2) is; $W = \{D_1, D_2, D_3, D_{x_1}, D_{x_2}, D_{x_3}\}$. Therefore, the availability is the sum of steady state probability of its working states, i.e.,

$$A_{\text{pump}} = P_{D_1} + P_{D_2} + P_{D_3} + P_{D_{x_1}} + P_{D_{x_2}} + P_{D_{x_3}}$$

The evaluated value of steady state availability is:

$$A_{\text{pump}} = 0.9971375783$$

The result obtained from the proposed model is in close agreement with the result (0.99970) following the same input data with different approach (Nel and Haarhoff 2011).

Step-V: optimization of CM interval using GA

The optimum CM interval to maximize the system availability is obtained by Genetic Algorithm approach—refer Sect. 4 for details. A code for availability following the above steps is written in Matlab and the availability function is called in the GA tool box of Matlab to find the optimum values of three CM intervals; CM intervals; T_1, T_2 and T_3, and system availability, A. The results obtained are: $T_1 = 9650$ h; $T_2 = 9600$ h; $T_3 = 9520$ h and $A = 0.999446816$.

The results obtained are realistic and attributed to the SMP model, incorporating multi-state degradation, appropriate distribution for modeling random failures, repairs and CM interval. The suggested methodology is useful for maintenance engineers in selecting suitable maintenance policies in the plant to achieve the system availability goals. Also, designers can use the methodology to design mechanical systems for desired availability by introducing the appropriate repair and level of repair.

Conclusion

In this paper, condition-based maintenance model of mechanical systems are proposed for steady-state availability analysis using the SMP. The system model is developed considering multi-state degradation, using condition-based maintenance. In addition to degradation, random failures are also considered. Weibull, log-normal, exponential and deterministic distributions are considered for degradation, repair, random failure and periodic CM, respectively. It is demonstrated through a case study that condition-monitoring can significantly improve the system availability. Also, the optimum CM interval is determined for maximizing the system availability using Genetic Algorithm approach.

The proposed methodology is quicker than the simulation-based approaches since results are obtained in a single run. Moreover, the results are more accurate than the results obtained from simulation approach as these are obtained from closed form solutions. Since appropriate distributions for the failure and the repair time are used, the results obtained with suggested methodology are more realistic. The model can be manipulated under different set of parameter values to see the impact of various maintenance policies on system availability. The proposed methodology is useful for maintenance engineers in decision making, and directly helps them for devising suitable maintenance/replacement policies in the plant.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1

See Table 4.

Table 4 Elements (non-zero) of global kernel matrix, $K(t)$—CBM model

k_{ij}	$k_{ij}(t)$	$k_{ij}(t)$ with specific distribution
$k_{D_1D_2}$	$\int_0^t \bar{F}_{D_1D_{x_1}} \bar{F}_{D_1F_r} dF_{D_1D_2}$	$\frac{\beta_{D_1D_2}}{\theta_{D_1D_2}^{\beta_{D_1D_2}}} \int_0^t U(x-T_1) x^{\beta_{D_1D_2}-1} e^{-[(\frac{x}{\theta_{D_1D_2}})^{\beta_{D_1D_2}} + \lambda_{D_1F_r} x]} dx$
$k_{D_1F_r}$	$\int_0^t \bar{F}_{D_1D_{x_1}} \bar{F}_{D_1D_2} dF_{D_1F_r}$	$\lambda_{D_1F_r} \int_0^t U(x-T_1) e^{-[(\frac{x}{\theta_{D_1D_2}})^{\beta_{D_1D_2}} + \lambda_{D_1F_r} x]} dx$
$k_{D_1D_{x_1}}$	$\int_0^t \bar{F}_{D_1F_r} \bar{F}_{D_1D_2} dF_{D_1D_{x_1}}$	$\int_0^t e^{-[(\frac{x}{\theta_{D_1D_2}})^{\beta_{D_1D_2}} + \lambda_{D_1F_r} x]} d(U(x-T_1))$
$k_{D_2D_3}$	$\int_0^t \bar{F}_{D_2D_{x_2}} \bar{F}_{D_2F_r} dF_{D_2D_3}$	$\frac{\beta_{D_2D_3}}{\theta_{D_2D_3}^{\beta_{D_2D_3}}} \int_0^t U(x-T_2) x^{\beta_{D_2D_3}-1} e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}} + \lambda_{D_2F_r} x]} dx$
$k_{D_2F_r}$	$\int_0^t \bar{F}_{D_2D_{x_2}} \bar{F}_{D_2D_3} dF_{D_2F_r}$	$\lambda_{D_2F_r} \int_0^t U(x-T_2) e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}} + \lambda_{D_2F_r} x]} dx$
$k_{D_2D_{x_2}}$	$\int_0^t \bar{F}_{D_2F_r} \bar{F}_{D_2D_3} dF_{D_2D_{x_2}}$	$\int_0^t e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}} + \lambda_{D_2F_r} x]} d(U(x-T_2))$
$k_{D_3F_r}$	$\int_0^t \bar{F}_{D_3D_{x_3}} dF_{D_3F_r}$	$\lambda_{D_3F_r} \int_0^t U(x-T_3) e^{-\lambda_{D_3F_r} x} dx$
$k_{D_3D_{x_3}}$	$\int_0^t \bar{F}_{D_3F_r} dF_{D_3D_{x_3}}$	$\int_0^t e^{-\lambda_{D_3F_r} x} d(U(x-T_3))$

Table 4 continued

k_{ij}	$k_{ij}(t)$	$k_{ij}(t)$ with specific distribution
$k_{F_rD_1}$	$F_{F_rD_1}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{F_rD_1}}{\sqrt{2}\sigma_{F_rD_1}}\right)$
$k_{D_{x_1}D_1}$	$p_{D_1}F_{D_{x_1}}$	$p_{D_1}(1-e^{-\lambda_{D_{x_1}}t})$
$k_{D_{x_1}mi_1}$	$(1-p_{D_1})F_{D_{x_1}}$	$(1-p_{D_1})(1-e^{-\lambda_{D_{x_1}}t})$
$k_{D_{x_2}D_2}$	$p_{D_2}F_{D_{x_2}}$	$p_{D_2}(1-e^{-\lambda_{D_{x_2}}t})$
$k_{D_{x_2}mi_2}$	$p_{mi_2}F_{D_{x_2}}$	$p_{mi_2}(1-e^{-\lambda_{D_{x_2}}t})$
$k_{D_{x_2}ma_2}$	$\{1-(p_{D_2}+p_{mi_2})\}F_{D_{x_2}}$	$\{1-(p_{D_2}+p_{mi_2})\}(1-e^{-\lambda_{D_{x_2}}t})$
$k_{D_{x_3}D_3}$	$p_{D_3}F_{D_{x_3}}$	$p_{D_3}(1-e^{-\lambda_{D_{x_3}}t})$
$k_{D_{x_3}mi_3}$	$p_{mi_3}F_{D_{x_3}}$	$p_{mi_3}(1-e^{-\lambda_{D_{x_3}}t})$
$k_{D_{x_3}im_3^1}$	$p_{im_3^1}F_{D_{x_3}}$	$p_{im_3^1}(1-e^{-\lambda_{D_{x_3}}t})$
$k_{D_{x_3}ma_3}$	$\{1-(p_{D_3}+p_{mi_3}+p_{im_3^1})F_{D_{x_3}}\}$	$\{1-(p_{D_3}+p_{mi_3}+p_{im_3^1})\}(1-e^{-\lambda_{D_{x_3}}t})$
$k_{mi_1D_1}$	$F_{mi_1D_1}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{mi_1D_1}}{\sqrt{2}\sigma_{mi_1D_1}}\right)$
$k_{mi_2D_2}$	$F_{mi_2D_2}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{mi_2D_2}}{\sqrt{2}\sigma_{mi_2D_2}}\right)$
$k_{mi_3D_3}$	$F_{mi_3D_3}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{mi_3D_3}}{\sqrt{2}\sigma_{mi_3D_3}}\right)$
$k_{im_3^1D_2}$	$F_{im_3^1D_2}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{mi_3^1D_2}}{\sqrt{2}\sigma_{mi_3^1D_2}}\right)$
$k_{ma_2D_1}$	$F_{ma_2D_1}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{ma_2D_1}}{\sqrt{2}\sigma_{ma_2D_1}}\right)$
$k_{ma_3D_1}$	$F_{ma_3D_1}(t)$	$\frac{1}{2}+\frac{1}{2}erf\left(\frac{\ln(x)-\mu_{ma_3D_1}}{\sqrt{2}\sigma_{ma_3D_1}}\right)$

Appendix 2

See Table 5.

Table 5 Non-zero elements of one step transition probability matrix, *Z*—CBM model

p_{ij}	Expressions of Elements of matrix, Z, $p_{ij}=k_{ij}(\infty)$	p_{ij} values
$p_{D_1D_2}$	$\frac{\beta_{D_1D_2}}{\theta_{D_1D_2}^{\beta_{D_1D_2}}}\int_0^{T_1}x^{\beta_{D_1D_2}-1}e^{-[(\frac{x}{\theta_{D_1D_2}})^{\beta_{D_1D_2}}+\lambda_{D_1F_r}x]}dx$	0.01047
$p_{D_1F_r}$	$\lambda_{D_1F_r}\int_0^{T_1}e^{-[(\frac{x}{\theta_{D_1D_2}})^{\beta_{D_1D_2}}+\lambda_{D_1F_r}x]}dx$	0.00793
$p_{D_1D_{x_1}}$	$1-\{k_{D_1D_2}(\infty)+k_{D_1F_r}(\infty)\}$	0.98160
$p_{D_2D_3}$	$\frac{\beta_{D_2D_3}}{\theta_{D_2D_3}^{\beta_{D_2D_3}}}\int_0^{T_2}x^{\beta_{D_2D_3}-1}e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}}+\lambda_{D_2F_r}x]}dx$	0.00187
$p_{D_2F_r}$	$\lambda_{D_2F_r}\int_0^{T_2}e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}}+\lambda_{D_2F_r}x]}dx$	0.00669
$p_{D_2D_{x_2}}$	$1-\{k_{D_2D_3}(\infty)+k_{D_2F_r}(\infty)\}$	0.99144
$p_{D_3F_r}$	$\lambda_{D_3F_r}\int_0^{T_3}e^{-\lambda_{D_3F_r}x}dx$	0.00479
$p_{D_3D_{x_3}}$	$1-k_{D_3F_r}(\infty)$	0.99521

Table 5 continued

p_{ij}	Expressions of Elements of matrix, Z, $p_{ij} = k_{ij}(\infty)$	p_{ij} values
$p_{D_{x_1}D_1}$	$p_{D_1}(1 - e^{-\lambda_{D_{x_1}} t})$	0.9
$p_{D_{x_1}mi_1}$	$(1 - p_{D_1})(1 - e^{-\lambda_{D_{x_1}} t})$	0.1
$p_{D_{x_2}ma_2}$	$1 - (p_{D_2} + p_{mi_2})$	0.3
$p_{D_{x_3}ma_3}$	$\{1 - (p_{D_3} + p_{mi_3} + p_{im_3^1})\}$	0.5

p_{ij}	Expressions of Elements of matrix, Z, $p_{ij} = k_{ij}(\infty)$	p_{ij} values	p_{ij}	Expressions of Elements of matrix, Z, $p_{ij} = k_{ij}(\infty)$	p_{ij} values
$p_{F_rD_1}$	1	1	$p_{mi_1D_1}$	1	1
$p_{D_{x_2}D_2}$	p_{D_2}	0.2	$p_{mi_2D_2}$	1	1
$p_{D_{x_2}mi_2}$	p_{mi_2}	0.5	$p_{mi_3D_3}$	1	1
$p_{D_{x_3}D_3}$	p_{D_3}	0.05	$p_{im_3^1D_2}$	1	1
$p_{D_{x_3}mi_3}$	p_{mi_3}	0.15	$p_{ma_2D_1}$	1	1
$p_{D_{x_3}im_3^1}$	$p_{im_3^1}$	0.3	$p_{ma_3D_1}$	1	1

Appendix 3

See Table 6.

Table 6 Mean sojourn time expressions—CBM model

τ_i	Sojourn time expressions	Sojourn time expressions with parameter values	Sojourn time values
τ_{D_1}	$\int_0^\infty \bar{F}_{D_1D_{x_1}}\bar{F}_{D_1F_r}\bar{F}_{D_1D_2}dx$	$\int_0^{T_2} e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}} + \lambda_{D_2F_r}x]}\mathrm{d}x$	237.84
τ_{D_2}	$\int_0^\infty \bar{F}_{D_2D_{x_2}}\bar{F}_{D_2F_r}\bar{F}_{D_2D_3}dx$	$\int_0^{T_2} e^{-[(\frac{x}{\theta_{D_2D_3}})^{\beta_{D_2D_3}} + \lambda_{D_2F_r}x]}\mathrm{d}x$	167.31
τ_{D_3}	$\int_0^\infty \bar{F}_{D_3D_{x_3}}\bar{F}_{D_3F_r}dx$	$\int_0^{T_3} e^{-\lambda_{D_3F_r}x}dx$	95.77
τ_{F_r}	$\int_0^\infty \bar{F}_{F_rD_1}dx$	$e^{\mu_{F_rD_1} + \frac{1}{2}\sigma^2_{F_rD_1}}$	8.22
$\tau_{D_{x_1}}$	$\int_0^\infty \bar{F}_{D_{x_1}}dx$	$\frac{1}{\lambda_{D_{x_1}}}$	1
$\tau_{D_{x_2}}$	$\int_0^\infty \bar{F}_{D_{x_2}}dx$	$\frac{1}{\lambda_{D_{x_2}}}$	2
$\tau_{D_{x_3}}$	$\int_0^\infty \bar{F}_{D_{x_3}}dx$	$\frac{1}{\lambda_{D_{x_3}}}$	3
τ_{mi_1}	$\int_0^\infty \bar{F}_{mi_1D_1}dx$	$e^{\mu_{mi_1D_1} + \frac{1}{2}\sigma^2_{mi_1D_1}}$	4.41
τ_{mi_2}	$\int_0^\infty \bar{F}_{mi_2D_2}dx$	$e^{\mu_{mi_2D_2} + \frac{1}{2}\sigma^2_{mi_2D_2}}$	6.89
τ_{mi_3}	$\int_0^\infty \bar{F}_{mi_3D_3}dx$	$e^{\mu_{mi_3D_3} + \frac{1}{2}\sigma^2_{mi_3D_3}}$	11.97
$\tau_{im_3^1}$	$\int_0^\infty \bar{F}_{im_3^1D_2}dx$	$e^{\mu_{im_3^1D_2} + \frac{1}{2}\sigma^2_{im_3^1D_2}}$	12.24
τ_{ma_2}	$\int_0^\infty \bar{F}_{ma_2D_1}dx$	$e^{\mu_{ma_2D_1} + \frac{1}{2}\sigma^2_{ma_2D_1}}$	9.53
τ_{ma_3}	$\int_0^\infty \bar{F}_{ma_3D_1}dx$	$e^{\mu_{ma_3D_1} + \frac{1}{2}\sigma^2_{ma_3D_1}}$	15.15

References

Ahmad R, Kamaruddin S, Azid I, Almanar I (2011) Maintenance management decision model for preventive maintenance strategy on production equipment. J Ind Eng Int 7(13):22–34

Ahmad A, Raissi S, Khalili-Damghani K (2017) A simulation-based optimization approach for free distributed repairable multi-state availability-redundancy allocation problems. Reliab Eng Syst Saf 157:177–191

Amari SV, McLaughlin L (2004) Optimal design of a condition-based maintenance model. In: Proceedings of IEEE annual reliability and maintainability symposium, Los Angeles, USA, 26–29 January, pp 528–533

Carr MJ, Wang W (2011) An approximate algorithm for prognostic modeling using condition monitoring information. Eur J Oper Res 211(1):90–96

Charles-Owaba OE, Oluleye AE, Oyawale FA, Oke SA (2008) An opportunity cost maintenance scheduling framework for a fleet of ships: a case study. J Ind Eng Int 4(6):64–77

Chen D, Trivedi KS (2005) Optimization for condition-based maintenance with semi-Markov decision process. Reliab Eng Syst Saf 90(1):25–29

Crocker J, Kumar UD (2000) Age-related maintenance versus reliability centered maintenance: a case study on aero-engines. Reliab Eng Syst Saf 67(2):113–118

Dennis A (2003) Application of MCS system reliability analysis. In: Proceedings of TAMU, 20th international pump users symposium. Department of Mechanical Engineering, Turbo machinery laboratory, Texas A & M University, College Station, Houston, Texas, USA, 17–20 March, pp 91–94

El-Damcese MA, Temraz NS (2011) Availability and reliability measures for multi-state system by using Markov reward model. Reliability Theory and Application 6(3):68–85

Fricks RM, Trivedi KS (1997) Modeling failure dependencies in reliability analysis using stochastic Petri nets. In: Proceedings of SCS Europe 11th European simulation multiconference, Istanbul, Turkey, 1–4 June, pp 1–22

Goldberg DE (2006) Genetic algorithm. Pearson Education India, New Delhi

Jager P, Bertsche B (2004) A new approach to gathering failure behaviour information about mechanical components based on expert knowledge. In: Proceedings of IEEE annual reliability and maintainability symposium, Los Angeles, USA, 26–29 January, pp 90–95

Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostic and prognostics implementing condition-based maintenance. Mech Syst Signal Process 20(7):1483–1510

Khanduja R, Tewari PC, Kumar D (2011) Steady state behavior and maintenance planning of bleaching system in a paper plant. J Ind Eng Int 7(12):39–44

Kulkrani VG (2009) Modeling and analysis of stochastic systems. Chapman and Hall, London

Kumar S, Tewari PC, Kumar S (2009) Performance evaluation and availability analysis of ammonia synthesis unit in a fertilizer plant. J Ind Eng Int 5(9):17–26

Kumar G, Jain V, Gandhi OP (2014) Steady-state availability analysis of repairable mechanical systems with opportunistic maintenance by using semi-Markov process. Int J Syst Assur Eng Manag 5(4):664–678

Kuo W, Rajendra VP, Tillman FA, Hwang CL (2001) Optimal reliability design fundamentals and applications. Cambridge University Press, London

Majid MAA, Nasir M (2011) Multi-state system availability model of electricity generation for a cogeneration district cooling plant. Asian J Appl Sci 4(4):431–438

MATLAB release 7.7 (2010) The Math Works Inc. Natick, Massachusetts

Nel DT, Haarhoff J (2011) Bulk water distribution power supply failures. J S Afr Inst Civ Eng 51(1):55–60

Pandian A, Ali A (2009) A review of recent trends in machine diagnosis and prognosis algorithm. In: Proceedings of IEEE World congress on nature & biologically inspired computing, Coimbatore, India, 9–11 December, pp 1731–1736

Qiang H, Zhihua D, Xiao Z (2010) Application of HSMM on NC machine's state recognition. In: Proceedings of IEEE international conference on e-health networking, digital ecosystems and technologies, Shenzhen, China, 17–18 April, pp 189–191

Sim SH, Endrenyi J (1988) Optimal preventive maintenance with repair. J IEEE Trans Reliab 37(1):92–96

Van PD, Voisin A, Levrat E, Iung B (2012) Condition-based maintenance with both perfect and imperfect maintenance actions. In: Proceedings of PHM society annual conference of the prognostics and health management society, Minnesota, USA, 23–27 September, pp 3–9

Wang W (2008) Condition based maintenance modeling. In: Kobbacy KAH, Murty DNP (eds) Complex system maintenance handbook, Part C. Springer, London, pp 111–131

Wang N, Sun S, Li S, Si S (2010) Modeling and optimization of deteriorating equipment with predictive maintenance and inspection. In: Proceedings of IEEE 17th international conference on industrial engineering and engineering management, Xiamen, China, 29–31 October, pp 942–946

Xie W, Hong Y, Trivedi K (2005) Analysis of a two-level software rejuvenation policy. Reliab Eng Syst Saf 87(1):13–22

Multidim Syst Sign Process
DOI 10.1007/s11045-017-0502-z

Color texture recognition by color information fusion using the non-extensive entropy

Seba Susan[1] · **Madasu Hanmandlu**[2]

Received: 8 December 2015 / Revised: 4 May 2017 / Accepted: 22 May 2017
© Springer Science+Business Media New York 2017

Abstract Color textures have a unique inter-relationship among its color planes since they contribute information about the same recurring pattern. The average information or entropy is thus presumed to be redundant across the color planes. This is the basis of our paper, which focuses on dimensionality reduction of color texture features by averaging the entropies across multidimensional color planes, while at the same time maintaining the high accuracy of color texture recognition. The mean operation was used in summarizing the original eleven-dimensional difference theoretic texture features for texture classification in Susan and Hanmandlu (IET Image Process 7(8):725–732, 2013). In this work, instead of the mean, we measure the entropy of the features across multidimensional color planes. The non-extensive entropy with the Gaussian information gain is used as the entropy measure for our experiments since it is non-linear and a good indicator of regular patterns in textures. Comparisons with the state-of-the-art prove the efficiency of our approach both in terms of accuracy and the reduced feature dimension.

Keywords Entropy · Color texture recognition · Difference theoretic texture features · Non-extensive entropy with Gaussian gain · Dimensionality reduction

1 Introduction

The application of entropy to pattern recognition problems is a popular trend since there is no measure better than entropy to gauge the uncertainty or chaos existing within a structured domain. The information from multiple domains can be integrated through the common platform of entropy. Entropy was one of the fourteen statistical texture descriptors given by Haralick et al. (1973). Its use as a single feature for texture classification was explored by the authors in Susan and Hanmandlu (2013a), in which the non-additive non-extensive

✉ Seba Susan
seba_406@yahoo.in

1 Department of Information Technology, Delhi Technological University, Delhi 110042, India

2 Electrical Engineering Department, Indian Institute of Technology, Delhi 110016, India

Published online: 02 June 2017

entropy with Gaussian gain was proposed as a single feature for the representation of correlated intensities within regular textures. The reason for its success is the non-linearity of its information gain which is a one-sided Gaussian function with zero mean. It can therefore represent correlations between image intensities with sparse probabilities very well. In this paper, we extend these advantages of the non-extensive entropy to color texture recognition by a two-step process: (1) find the entropy of the local texture cues defined at pixel level in each of the color planes (2) explore the redundancy of the computed entropies across the three color planes for dimensionality reduction. The texture cues used in our work are the difference theoretic texture features (DTTF) proposed by the authors in Susan and Hanmandlu (2013b), primarily due to its advantage of being pixel defined in the initial stage, which is then subsequently summarized by averaging over all the pixel positions. These cues are eleven in number and demonstrate highly efficient texture recognition as compared to the LBP and LBPV operators. The concatenated feature dimension for color textures was only 33 as compared to 177 for LBPV features (Susan and Hanmandlu 2013b). Most of the works in literature for color-texture recognition (also termed the parallel methods), involve concatenation of the texture features in R, G, B color domains for classification (Ilgner et al. 2003; Messer and Kittler 1999; Porebski et al. 2008; Zhu et al. 2010). In Porebski et al. (2008) Haralick's fourteen features are computed from the co-occurrence matrix of the LBP texture features. Dimensionality reduction is usually performed by separate feature ranking and selection techniques like NMIFS (Estevez et al. 2009) and the maximum fuzzy entropy (Luukka 2011) technique. In our paper, we incorporate dimensionality reduction at the feature extraction stage itself, by utilizing the redundancy of entropies of local texture cues across the three color planes. Recent works such as Alvarez et al. (2012) advocate the use of sparse color texture descriptors using texton theory in a 3D texture space. An integrative co-occurrence matrix approach to color texture recognition was given in Palm (2004) where Kolmogorov distance was applied to the multi-channel matrices. In Alvarez and Vanrell (2012) the co-joint texton descriptors (JTD) that use the co-occurrence of shape and color attributes to represent color textures were presented. Recent research in color texture recognition combines different color-texture features and applies machine learning on color planes for dimensionality reduction (Khan et al. 2015). Out of all the single methods tested in Khan et al. (2015), the completed local binary pattern (CLBP) (Guo et al. 2010b) was proved to be best. Local ternary patterns (LTPs) (Tan and Triggs 2010) introduce a three-value alternative to LBP that is proved to be more robust and efficient. The combination of texture primitive pattern units and statistical features in Ershad (2011) was found effective for color texture classification. In this paper, we use the co-occurrence probabilities of the DTTF defined at each pixel position to obtain the entropy value in a 2D color plane. The entropy values across multiple color planes is then averaged to yield the final feature vector.

The organization of this paper is as follows: Sect. 2 contains a review of related texture cues, particularly the DTTF and the non-extensive entropy with Gaussian gain. Section 3 gives the steps for the proposed color-texture recognition, Sect. 4 discusses the experimentation and the results and Sect. 5 draws the final conclusions.

2 A review of the related works

2.1 Difference theoretic texture features (DTTF)

A new set of features called the DTTF, based on local and global gray-level differences, that outperformed LBP (Guo et al. 2010a; Ojala et al. 2002a) and the MRS4 filter bank

(Leung and Malik 2001) in its classification performance (98.8457% on 64,800 samples of twenty five monochrome Brodatz textures), was proposed by the authors in Susan and Hanmandlu (2013b) for high accuracy texture classification. The utility of this feature set for the representation of color textures was also proved in Susan and Hanmandlu (2013b) for fifteen color textures from the *Outex* dataset, by concatenating the DTTF extracted from the R, G, B color planes respectively. A high accuracy of 99.89% was achieved in the experiments performed. It is specially noted here, that the eleven-dimension feature set is computed locally for each pixel and subsequently averaged over all the pixels in the texture image to generate the feature vector for recognition.

In our work, we replace the average by the entropy value with an aim to fuse the average color information from different color planes and hence reduce the feature dimension. The formulation of the original DTTF is briefly described below.

The DTTF comprises of the local and global differences and their individual and joint probabilities of occurrence in the image. The local differences were computed between the pixel and its immediate neighbor in the horizontal (H), vertical (V) and diagonal (D) directions. The global difference was obtained by subtracting the pixel intensity from the mean luminance Y of the image. The eleven-dimensional feature set **D** obtained by averaging the features over all pixels is given by,

$$\begin{aligned}\mathbf{D} = [&absdiffH, absdiffV, absdiffD, pdiffH, pdiffV,\\ &pdiffD, absY, pY, pdiffHY, pdiffVY, pdiffDY]\end{aligned} \tag{1}$$

In (1) the terms *absdiffH*, *absdiffV*, *absdiffD* and *absY* refer to the average of the local (horizontal, vertical and diagonal) and the global differences pertaining to each pixel in the image. The histograms of the local and global differences are plotted and the probability values *pdiffH*, *pdiffV*, *pdiffD*, *pY* are mapped to each pixel position based on closest bin-matching with the pixel intensity. The joint histograms are then plotted and the probability values *pdiffHY*, *pdiffVY*, *pdiffDY* are also mapped back to each pixel position. For further reference the reader is referred to Susan and Hanmandlu (2013b) where all the equations are described in detail.

2.2 The non-extensive entropy with Gaussian gain

In order to increase the non-linearity of the exponential family of entropies, the non-extensive entropy function was proposed in Susan and Hanmandlu (2013a), which has an exponential information gain with a quadratic exponent. The non-extensive entropy with Gaussian gain was proved successful in the representation of correlated texture patterns containing non-additive information (98.3% for fifteen Brodatz textures), and it outperformed the Shannon, Renyi, Tsallis, Pal and Pal entropies (Shannon 1948; Renyi 1961; Tsallis et al. 2001; Pal and Sankar 1991) and the MR4 filter based technique (Varma and Zisserman 2005) in the experiments performed.

Consider a random variable $X = \{x_1, x_2, \ldots, x_n\}$ with the associated probabilities $P = \{p_1, p_2, \ldots, p_n\}$. Assume that the probability distribution is complete i.e. $p_i \in [0, 1]$, $\sum_{i=1}^{n} p_i = 1$, where n is the number of probabilistic experiments. Let the information gain on the ith event of X with an associated probability p_i be defined by the Gaussian function $e^{-p_i^2}$. The non-extensive entropy with Gaussian gain is thus defined as

$$H(P) = E(I(p_i)) = \sum_{i=1}^{n} p_i I(p_i) = \sum_{i=1}^{n} p_i e^{-p_i^2} \tag{2}$$

The normalized form of the entropy in (2) is

$$H_N = \frac{(H - H_{\min})}{(H_{\max} - H_{\min})} \tag{3}$$

$$\text{i.e. } H_N = \frac{(H - e^{-1})}{e^{-\frac{1}{n^2}} - e^{-1}} \tag{4}$$

The non-extensive entropy in (2) being non-additive, is a good identifier of the correlated intensities in texture images that translate to higher co-occurrence probabilities, in contrast to the additive entropies like Shannon entropy that are more suited for chaotic and non-ordered domains. Since most natural textures contain some degree of regularity of structure, the experiments in Susan and Hanmandlu (2013a) showed the best results for the non-extensive entropy.

2.3 Some other related texture cues

A brief description of the completed local binary pattern (CLBP) (Guo et al. 2010b) and the local ternary pattern (LTP) (Tan and Triggs 2010) texture cues are included here for reference, since they too involve the computation of histograms of gray-level differences in local neighborhoods which is an underlying concept in our work. CLBP proposed by Guo et al. constitutes of binary encoding of the gray-level value of the center pixel in a neighborhood and the sign and magnitude of the local gray-level differences. The *CLBP_Center*, *CLBP_Sign* and *CLBP_Magnitude* encoding schemes are used respectively. Once the encoding is complete, global histograms are computed for the local binary patterns created. CLBP thus scores over its predecessor, the local binary pattern (LBP) that encodes only the sign of these local gray-level differences. The LTP operator proposed by Tan and Triggs (2010) is

1. *Input the RGB color texture image*
2. *For each texture sample resized to 128x128 do*
   - i. *For each of the R, G, B color planes (for RGB color model), find the Difference Theoretic Texture Features defined at each pixel by skipping the step of averaging over MxN pixels that yields the one-dimensional final feature vector in [3]. The 11-dimension feature set is now defined for each pixel (i,j) in the image, for each of the color planes.*
   - ii. *Find the Non-Extensive entropy using (2) for each of the eleven 2D DTTF feature matrices using the gray-level co-occurrence probabilities computed from each matrix. The offset distance is set to 31 for each of the (128x128) 2D feature matrices (since we assume as in [2] that each quarter of the 128x128 texture image encloses one or more repetitions of the texture pattern), and the quantization levels for the computation of co-occurrence probabilities are set to 16. The settings are defined after lots of experimental trials for the best results.*
   - iii. *Concatenate the entropy values of the eleven feature matrices from each color plane to form the 33-dimension concatenated entropy feature vector for comparison purpose.*
   - iv. *Average the entropies for each of the 11 features across the three color channels, due to redundancy in entropy values across the color planes. Apply the (reduced) 11 dimensional feature set for classification.*

Fig. 1 The methodology for the proposed color-texture recognition

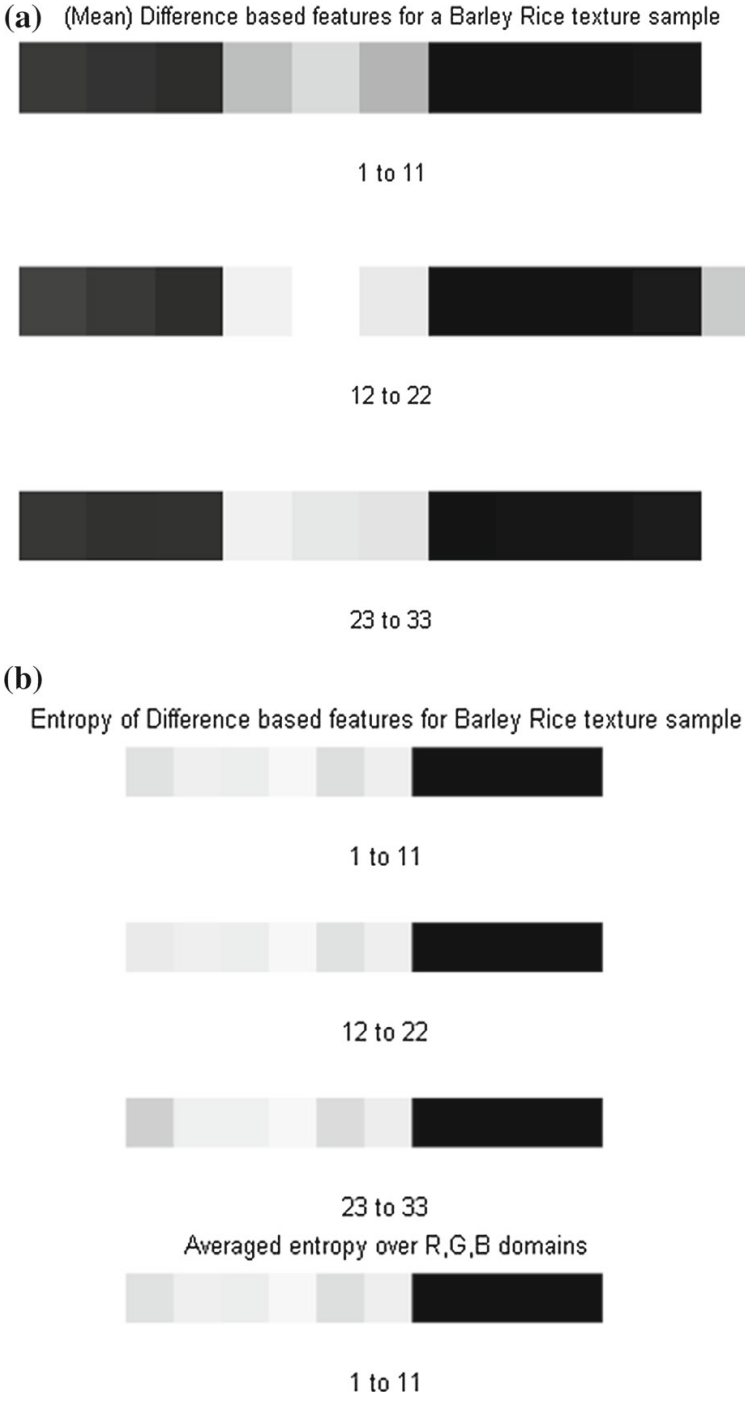


Fig. 2 The difference theoretic texture feature set and the non-extensive entropy of difference theoretic texture features set for *Barley Rice* texture sample for the red (1–11), green (12–22) and blue (23–33) color domains

a three-valued code [−1, 0, +1], as opposed to LBP's two-valued code [1, 0]. In each local neighborhood, upper and the lower binary patterns are created and concatenated to form the LTP feature vector. The accuracy of CLBP is reported to be higher, due to the hybrid distributions used. The LTP, however, is more robust to noise. In both methods, the local information in the form of binary patterns are compiled in the form of histograms that in turn form the feature vector for recognition.

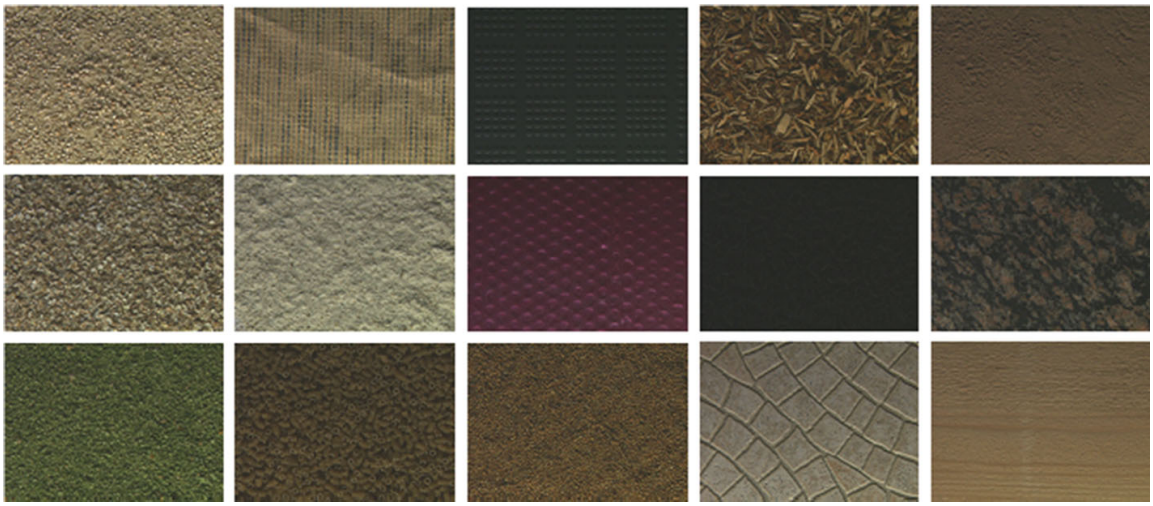


Fig. 3 Fifteen color textures from the *Outex_TC_00000* database (Ojala et al. 2002b), from *top to bottom* and *left to right*: *Barley Rice, Canvas, Carpet, Chips, Crushed Stone, Flakes, Flour, Foam, Fur, Granite, Groats, Pasta, Seeds, Tile, Wood*

3 The proposed methodology for color texture recognition using entropy of color texture cues

It is assumed in our work that the entropies of texture features are redundant across the R,G,B color planes since all color planes though visually different contribute the same information. The entropies of the pixel-level defined texture cues are redundant across color planes, and the average of the three color entropies would yield a new fused feature vector. DTTF is one such feature set where the features are locally derived at each pixel position and subsequently averaged over all the pixels to yield a one-dimensional feature vector. So instead of averaging, we propose the entropy measurement as a more useful alternative for color textures.

Thus the primary contributions highlighted of the work done in this paper are:

1. A novel method of color information fusion by averaging the entropies of pixel defined features across the three color planes.
2. Dimensionality reduction since the feature size is reduced to one-third of the concatenated feature-size for the color planes.
3. The reduced feature vector gives a general idea of the average information content in the image measured in terms of the average non-extensive entropy. Thus, it helps in identifying the salient object in the image in case of natural color images.

The proposed texture recognition incorporating entropy redundancy for color information fusion and feature dimensionality reduction is outlined in Fig. 1.

Figure 2 shows the 33 concatenated (mean) DTTF and the entropy of DTTF for a Barley Rice texture sample, with images scaled in the range of 0–1. The redundancy in the entropy values across R, G, B color domains is clearly visible from Fig. 2b but not from Fig. 2a indicating that the (mean) DTTF are different for the different color planes.

3.1 Addressing problems of affine-invariance, computational complexity and noise

Since DTTF is affine-invariant i.e. the features are invariant to changes in rotation, illumination and scale of the texture image, therefore the non-extensive entropy of these features inherit the affine-invariant properties of DTTF. The computation of the DTTF hardly takes

Table 1 Color texture recognition results (in %) using the SVM classifier for the fifteen categories of the *Outex_00000* dataset (Validation)

Texture category	Difference theoretic texture features set (Susan and Hanmandlu 2013b)	Proposed non-extensive entropy of color difference theoretic texture features set		LBPV (Guo et al. 2010a)	LBPV (Guo et al. 2010a)	JTD (Alvarez and Vanrell 2012)	CLBP (Guo et al. 2010b)	LTP (Tan and Triggs 2010)	Texture primitive unit (Ershad 2011)
	(33 features)	(33 feat.)	(11 feat.)	(177 features)	(30 features)	(15,625 features)	(600 features)	(768 features)	(12 features)
Barley Rice	100	100	100	100	100	100	100	99.17	83.33
Canvas	100	100	100	100	100	100	100	100	97.08
Carpet	100	100	99.17	100	93.3	100	100	100	97.08
Chips	100	99.17	98.33	100	99.17	100	100	100	96.67
Crushed Stone	100	100	100	100	100	100	100	100	96.67
Flakes	100	100	100	100	100	100	99.17	100	91.67
Flour	100	100	100	99.17	100	100	100	100	95.42
Foam	100	100	99.17	100	98.33	100	100	100	95.42
Fur	100	100	100	100	99.17	100	100	100	97.92
Granite	100	100	100	98.33	93.3	98.33	98.33	99.17	97.92
Groats	100	100	100	100	100	100	100	100	95
Pasta	100	100	100	98.33	99.17	99.17	100	99.17	95.83
Seeds	100	100	100	100	100	100	100	100	95.42
Tile	100	100	100	99.17	99.17	100	100	100	97.92
Wood	98.33	98.33	99.17	100	100	99.17	100	100	90
Average	99.89	99.83	99.72	99.67	98.77	99.77	99.83	99.83	94.89

Table 2 Color texture recognition results (in %) using the SVM classifier for the fifteen categories of the *Outex_00000* dataset (cross-validation)

Texture category	Difference theoretic texture features set (Susan and Hanmandlu 2013b)	Proposed non-extensive entropy of color difference theoretic texture features set		LBPV (Guo et al. 2010a)	LBPV (Guo et al. 2010a)	JTD (Alvarez and Vanrell 2012)	CLBP (Guo et al. 2010b)	LTP (Tan and Triggs 2010)	Texture primitive unit (Ershad 2011)
	(33 features)	(33feat.)	(11feat.)	(177 features)	(30 features)	(15,625 features)	(600 features)	(768 features)	(12 features)
Barley Rice	100	100	100	99.17	98.33	99.17	100	100	87.08
Canvas	100	100	99.17	100	100	99.17	100	100	97.08
Carpet	100	100	99.17	100	93.33	100	100	100	95.83
Chips	100	100	99.17	99.17	100	100	100	100	96.25
Crushed Stone	99.17	100	100	100	100	100	100	100	97.92
Flakes	99.17	99.17	99.17	100	99.17	100	100	99.17	91.25
Flour	100	99.17	99.17	100	98.33	100	100	100	97.50
Foam	100	100	99.17	99.17	93.33	100	100	100	95
Fur	100	100	100	99.17	96.67	100	100	100	97.92
Granite	100	100	100	98.33	98.33	98.33	100	98.33	97.92
Groats	100	100	99.17	100	98.33	100	100	100	95.42
Pasta	100	100	100	99.17	99.17	99.17	99.17	100	97.08
Seeds	100	100	100	100	100	100	100	100	95
Tile	100	100	100	100	98.33	100	100	100	97.92
Wood	100	100	100	100	99.17	99.17	100	100	90.42
Average	99.89	99.89	99.61	99.61	98.16	99.66	99.94	99.83	95.31

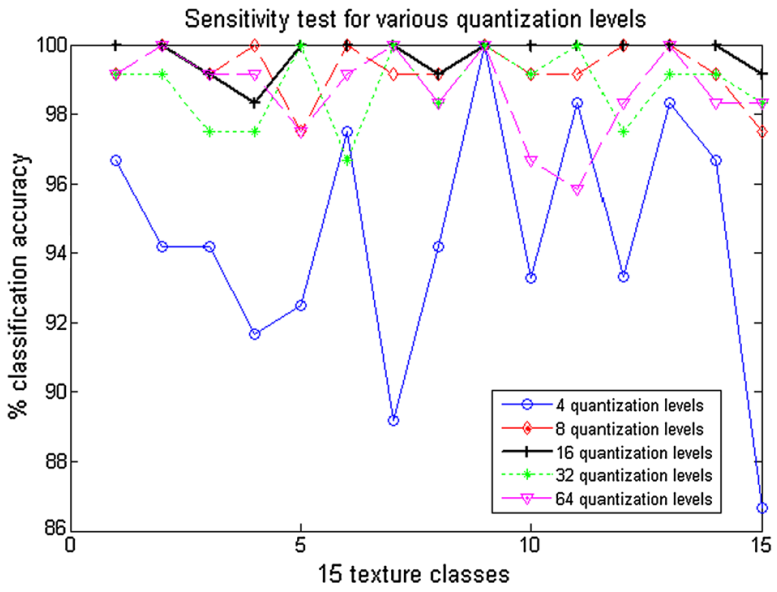


Fig. 4 The comparison of various quantization levels for the fifteen classes of textures of *Outex_TC_00000*

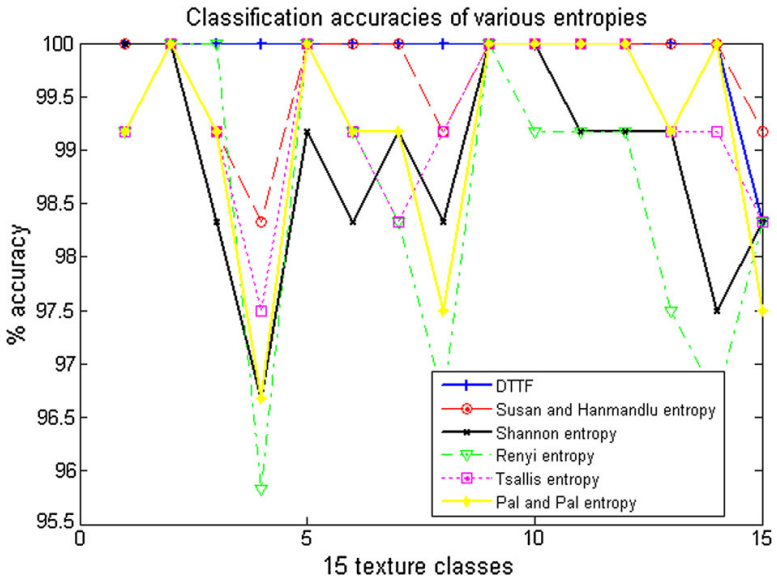


Fig. 5 The comparison of various entropies for the fifteen classes of textures of *Outex_TC_00000*

a minute on a 2.6 GHz PC and the increase in computation due to the calculation of entropy instead of the mean is very marginal.

For additive white noise, the effect on the image would be similar to the effect of constant changes in illumination which is addressed to by DTTF. However for non-linear sources of noise such as salt and pepper noise, the results would be impacted and so median filtering is added as a requisite for image pre-processing in our method.

4 Experimental results and discussions

The color-texture recognition experiments are conducted on a 2.6 GHz PC and the software implementation is done on MATLAB 7.9 version platform for the RGB, HSV, CIELAB and YCbCr color spaces on a variety of color-texture datasets. In the first stage of experiments,

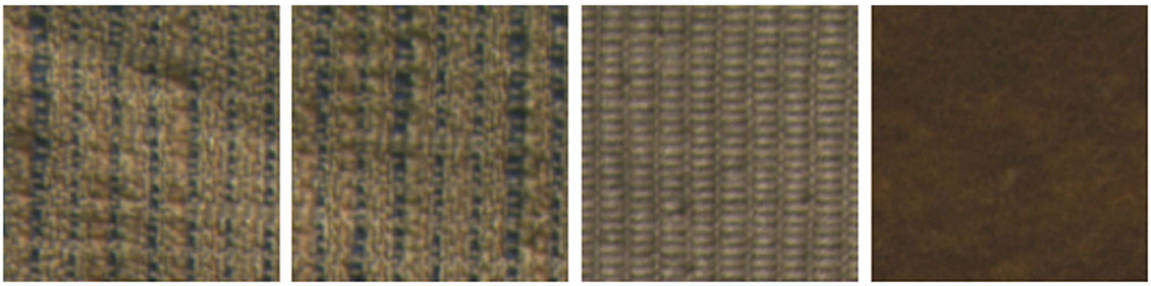


Fig. 6 Some sample images from the *Outex_TC_00013* database

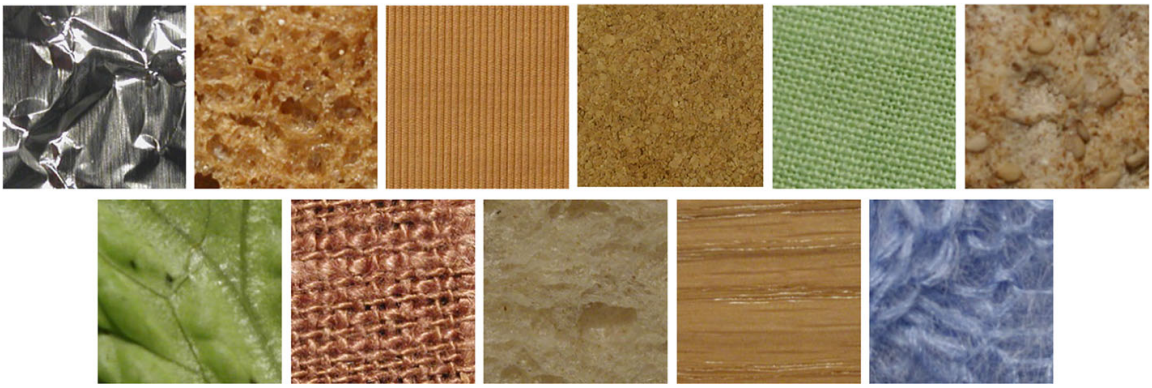


Fig. 7 Eleven color texture categories from the KTH-TIPS-2a database (http://www.nada.kth.se/cvap/databases/kth-tips/index.html), from top to bottom and left to right: *aluminum foil, brown bread, corduroy, cork, cotton, cracker, lettuce leaf, linen, white bread, wood, wool*

each texture image of dimension 746 × 538 belonging to one of fifteen categories (*Barley Rice, Canvas, Carpet, Chips, Crushed Stone, Flakes, Flour, Foam, Fur, Granite, Groats, Pasta, Seeds, Tile, Wood* : samples shown in Fig. 3) of the *Outex_00000* database Ojala et al. (2002b) is divided into 16 non-overlapping samples of size 128 × 128. Since there are fifteen texture classes in all, (15 × 16 =)240 color texture samples are derived out of which 120 are for used for training and 120 for testing. The SVM classifier is used for the classification, the settings of which are the same as in Susan and Hanmandlu (2013b). The binary SVM classifier is used for our multi-class classification experiments by using the one-against-all approach. The binary classification accuracies are averaged over all the classes to yield the final score. The eleven dimensional DTTF is extracted from each color channel as for a gray level intensity image, and then concatenated in the order of R,G,B for the RGB color model, H,S,V for the HSV color model, L,A,B for the CIELAB color model and Y,Cb,Cr for the YCbCr color model. Tables 1 and 2 list the results of color texture recognition for the *Outex_00000* dataset together with the comparison with state-of-the-art techniques namely, the Local Binary Pattern with Variance (LBPV) with 30 and 177 features (Guo et al. 2010a), an integrative approach- the joint texton descriptor (JTD) (Alvarez and Vanrell 2012), completed local binary pattern (CLBP) (Guo et al. 2010b), local ternary pattern (LTP) (Tan and Triggs 2010) and texture primitive units (TPUs) with statistical features (Ershad 2011). The 11-dimensional reduced dimension entropy feature set gives a high accuracy of 99.72% albeit lower than the concatenated entropy features. As observed from Table 1 the proposed reduced-dimension entropy feature set at 99.72% (11 features) performs better than the LBPV (177 features—99.67% and 30 features—98.77%) (Guo et al. 2010a) and TPU (Ershad 2011) at 94.89%, and is comparable to the JTD color-texture features (15,625 features) (Alvarez and Vanrell 2012) at 99.77%. The 33-dimensional concatenated Non-Extensive entropies

Table 3 The AUC of the receiver operating characteristic (ROC) curves for the 68 categories of the *Outex_TC_00013* dataset for both validation (V) and cross-validation (CV) using the SVM classifier

Methods	Number of features	RGB color model		HSV color model		CIELAB color model		YCbCr color model	
		V	CV	V	CV	V	CV	V	CV
Difference theoretic texture features set (Susan and Hanmandlu 2013b)	33	0.983	0.979	0.986	0.984	0.980	0.977	0.968	0.966
Non-extensive entropy of color difference theoretic texture features set	33	0.935	0.924	0.879	0.878	0.915	0.905	0.768	0.780
Proposed averaged non-extensive entropy of color difference theoretic texture features set	11	0.974	0.963	0.965	0.968	0.981	0.979	0.954	0.948
LBPV (Guo et al. 2010a)	177	0.958	0.956	0.739	0.731	0.942	0.939	0.938	0.941
LBPV (Guo et al. 2010a)	30	0.931	0.923	0.721	0.710	0.909	0.909	0.928	0.905
JTD (Alvarez and Vanrell 2012)	15,625	0.991	0.987	0.989	0.984	0.937	0.926	0.877	0.890
CLBP (Guo et al. 2010b)	600	0.9299	0.9320	0.9299	0.9320	0.8969	0.9042	0.9341	0.9355
LTP (Tan and Triggs 2010)	768	0.9457	0.9397	0.9457	0.9397	0.50	0.50	0.50	0.50
Texture primitive unit (Ershad 2011)	12	0.9336	0.963	0.9452	0.9643	0.8863	0.9492	0.8497	0.945

Table 4 The classification accuracies (%) for the color textures in Fig. 7 belonging to KTH-TIPS-2a database (http://www.nada.kth.se/cvap/databases/kth-tips/index.html) for the RGB and CIELAB color spaces using several classifiers (cross-validation results shown in parenthesis in each cell) (mean ± standard deviation of accuraces)

Classifier	Difference theoretic texture features set (Susan and Hanmandlu 2013b)	Proposed non-extensive entropy of color difference theoretic texture features set		LBPV (Guo et al. 2010a)	LBPV (Guo et al. 2010a)	JTD (Alvarez and Vanrell 2012)	CLBP (Guo et al. 2010b)	LTP (Tan and Triggs 2010)	Texture Primitive Unit (Ershad 2011)
(Color model)	(33 features)	(33 feat.)	(11feat.)	(177 features)	(30 features)	(15,625 features)	(600 features)	(768 features)	(12 features)
SVM (RGB)	95.53 ± 2.4	94.79 ± 1.9	94.03 ± 2.0	93.72 ± 2.8	93.28 ± 2.6	96.94 ± 1.9	93.58 ± 2.7	90.91 ± 0.0	84.85 ± 0.060
	(95.06 ± 2.3)	(94.16 ± 2.4)	(93.81 ± 2.1)	(93.53 ± 2.7)	(92.88 ± 2.2)	(96.91 ± 1.6)	(93.31 ± 3.0)	(90.91 ± 0.0)	(83.79 ± 0.072)
SVM (CIELAB)	95.82 ± 2.0	96.48 ± 1.5	95.61 ± 1.5	94.68 ± 2.8	93.83 ± 2.4	95.17 ± 1.7	91.24 ± 0.8	90.91 ± 0.0	83.57 ± 0.085
	(95.60 ± 2.0)	(96.43 ± 1.6)	(95.68 ± 2.0)	(94.32 ± 2.5)	(93.84 ± 2.2)	(94.70 ± 1.6)	(91.24 ± 0.7)	(90.91 ± 0.0)	(82.39 ± 0.079)
Bagged tree ensemble (RGB)	84.22 ± 10	81.73 ± 11.7	73.61 ± 15.7	87.48 ± 5.4	81.92 ± 8.8	64.93 ± 31.9	85.94 ± 0.8	9.09 ± 30	69.44 ± 0.0
	(80.24 ± 8)	(78.53 ± 9.9)	(69.52 ± 14.0)	(84.1 ± 6.8)	(78.94 ± 13.3)	(57.48 ± 30.8)	(84.90 ± 0.6)	(9.09 ± 30)	(69.27 ± 0.0)
Bagged tree ensemble (CIELAB)	86.34 ± 8.3	85.66 ± 7.4	79.80 ± 10.1	45.56 ± 33	84.05 ± 7.7	45.56 ± 33.0	89.21 ± 4.3	9.09 ± 30	67.36 ± 0.0
	(83.49 ± 10.5)	(81.77 ± 11.9)	(74.48 ± 12.2)	(34.65 ± 28)	(81.75 ± 8.6)	(37.65 ± 28.35)	(84.29 ± 9.9)	(9.09 ± 30)	(67.36 ± 0.0)



Fig. 8 Natural images from the Caltech-256 database (Griffin et al. 2007) belonging to the categories *butterflies* and *octopus*

give a higher recognition rate of 99.83% quite near to the performance of the concatenated 33-dimensional mean DTTF (99.89%). The comparison results confirm the efficiency of our method with high percentage accuracies obtained at the lowest possible feature dimension for color textures. A sensitivity test is carried out to find the optimal number of quantization levels for entropy calculation, the results of which are shown in Fig. 4. The classification accuracy for 16 quantization levels is verified to be the highest at 99.72%, better than 4 (94.4%), 8 (99.28%), 32 (98.78%), 64 (98.67%) quantization levels. The results of our experiments on different entropy functions, shown in Fig. 5 gave the best results for the non-extensive entropy with Gaussian gain (99.72%) proposed by Susan and Hanmandlu (2013a), followed by Tsallis entropy (99.27%) and Pal and Pal entropy (99.17%), while the additive entropies of Shannon (98.89%) and Renyi (98.61%) do not perform so well. We also perform the experimentation on the complex dataset *Outex_TC_00013* containing 68 color texture classes (examples shown in Fig. 6) and also on the KTH-TIPS-2a dataset (http://www.nada.kth.se/cvap/databases/kth-tips/index.html) containing eleven categories of color textures shown in Fig. 7. The results summarized in Table 3 for 1360 images of the *Outex_TC_00013* dataset each of size 128x128, show a consistently high performance for the proposed method, for all the color models, though the CIELAB color space gives the best results. On the other hand JTD gives good results in the RGB color space. The results on the KTH-TIPS-2a dataset are summarized in Table 4. We compare the SVM results with the bagged tree ensemble classifier (trees $= 100$), and we find the SVM classifier, being highly non-linear, more suited for testing the entropy values. We also test our method on some natural color images belonging to the *Butterflies* and *Octopus* object categories from the Caltech-256 dataset (Griffin et al. 2007) shown in Fig. 8. The percentage of classification accuracies are summarized in Table 5 for the natural images in Fig. 8 and the results indicate the highest accuracy for the proposed method in the CIELAB color space for both validation and cross-validation, while the DTTF performs very well in the RGB color space (however for validation only). The JTD on the other hand gives a reasonably good performance in the RGB, CIELAB color spaces for both validation and cross-validation. This implies that in the case of natural images, integrative approaches (like the proposed averaging of non-extensive entropies) are more successful than concatenated features since they do not depend on the actual value of colors in the scene but focus on the fusion of cumulative information from all color channels.

Table 5 The classification accuracies (%) using SVM classifier for the natural images in Fig. 8 belonging to two classes for the RGB and CIELAB color spaces using the SVM classifier

Methods	Number of features	RGB color model		HSV color model		CIELAB color model		YCbCr color model	
		V	CV	V	CV	V	CV	V	CV
Difference theoretic texture features set (Susan and Hanmandlu 2013b)	33	**80**	55	60	55	60	60	45	40
Non-extensive entropy of color difference theoretic texture features set	33	50	50	30	40	60	50	60	55
Proposed averaged non-extensive entropy of color difference theoretic texture features set	11	45	65	65	50	**75**	**80**	40	60
LBPV (Guo et al. 2010a)	177	50	50	45	50	45	40	50	60
LBPV (Guo et al. 2010a)	30	50	50	50	50	50	50	50	55
JTD (Alvarez and Vanrell 2012)	15,625	**70**	65	55	60	**70**	55	60	65
CLBP (Guo et al. 2010b)	600	50	50	50	50	50	50	50	50
LTP (Tan and Triggs 2010)	768	50	50	50	50	50	50	50	50
Texture primitive units (Ershad 2011)	12	50	50	50	50	50	50	50	50

Values greater than 70% are highlighted in bold

5 Conclusions

A new color texture recognition scheme is introduced in this work based on the multidimensional color entropy of the DTTF set. The redundancy of entropy across multiple color planes is explored for the dimensionality reduction. The proposed color information fusion using the non-extensive entropy with Gaussian gain can be extended to other content-based image analysis problems where the actual color is not significant.

References

Alvarez, S., Salvatella, A., Vanrell, M., & Otazu, X. (2012). Low dimensional and comprehensive color texture recognition. *Computer Vision and Image Understanding*, *116*, 54–67.

Alvarez, S., & Vanrell, M. (2012). Texton theory revisited: A bag-of-words approach to combine textons. *Pattern Recognition*, *45*(12), 4312–4325.

Ershad, S. F. (2011). *Color texture classification approach based on combination of primitive pattern units and statistical features*. arXiv preprint arXiv:1109.1133.

Estevez, P. A., Tesmer, M., Perez, C. A., & Zurada, J. M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks*, *20*(2), 189–201.

Griffin, G., Holub, A., & Perona, P. (2007). *Caltech-256 object category dataset*.

Guo, Z., Zhang, L., & Zhang, D. (2010b). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing*, *19*(6), 1657–1663.

Guo, Z., Zhang, L., & Zhang, D. (2010a). Rotation invariant texture classification using LBP Variance with global matching. *Pattern Recognition*, *43*, 706–719.

Haralick, R. M., Shanmugan, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*, 610–621.

Ilgner, J. F. R., Palm, C., SchSutz, A. G., Spitzer, K., Westhofen, M., & Lehmann, T. M. (2003). Colour texture analysis for quantitative laryngoscopy. *Acta Otorrinolaringologica*, *123*, 730–734.

Khan, Fahad Shahbaz, Anwer, Rao Muhammad, Weijer, Joost van de, Felsberg, Michael, & Laaksonen, Jorma. (2015). Compact color-texture description for texture classssification. *Pattern Recognition Letters*, *51*, 16–22.

Leung, T., & Malik, J. (2001). Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, *43*(1), 29–44.

Luukka, P. (2011). Feature selection using fuzzy entropy measures with similarity classifier. *Expert Systems with Applications*, *38*, 4600–4607.

Messer, K., & Kittler, J. (1999). A region-based image database system using colour and texture. *Pattern Recognition Letters*, *20*, 1323–1330.

Ojala, T., Maenpaa, T., Pietikainen, M., Viertola, J., Kyllonen, J., & Huovinen, S. (2002b). Outex-new framework for empirical evaluation of texture analysis algorithms. In *Proceedings of the 16th international conference on pattern recognition, 2002* (Vol. 1, pp. 701–706). IEEE.

Ojala, T., Pietikaenen, M., & Maenepae, T. (2002a). Multi-resolution gray scale and rotation invariant texture classification with LBP. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971–987.

Palm, C. (2004). Color texture classification by integrative co-occurence matrices. *Pattern Recognition*, *37*, 965–976.

Pal, N., & Pal, S. (1991). Entropy—A new definition and its applications. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(5), 1260–1270.

Porebski, A., Vandenbrouke, N., & Macaire, L. (2008). Haralick feature extraction from LBP images for color texture classification. In *Proceedings of IEEE conference on image processing theory, tools and applications*.

Renyi, A. (1961). On measures of entropy and information. *Fourth Berkeley Symposium*, *1*, 547–561.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system Technical Journal*, *27*, 379–423.

Susan, S., & Hanmandlu, M. (2013a). A non-extensive entropy feature and its application to texture classification. *Neurocomputing*, *120*, 214–225.

Susan, S., & Hanmandlu, M. (2013b). Difference theoretic feature set for scale-, illumination-and rotation-invariant texture classification. *IET Image Processing*, *7*(8), 725–732.

Tan, Xiaoyang, & Triggs, Bill. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, *19*(6), 1635–1650.

Tsallis, C., Abe, S., & Okamoto, Y. (2001). *Non extensive statistical mechanics and its applications. Lecture notes in physics*. Berlin: Springer.

Varma, M., & Zisserman, A. (2005). A statistical approach to texture classification using single images. *International Journal of Computer Vision*, *62*, 61–81.

Zhu, C., Bichot, C.-E., & Chen, L. (2010). Multi-scale color local binary patterns for visual object classes recognition. In *ICPR 2010* (pp. 3065–3068).



Seba Susan She received her Ph.D in Image Processing from the Electrical Engineering Department of Indian Institute of Technology (IIT), Delhi in 2014. She is currently Assistant Professor in the Department of Information Technology, Delhi Technological University (DTU), Delhi. Her research areas are statistical pattern recognition with machine learning, soft computing and image processing.



Madasu Hanmandlu (SM'06) received his B.E. degree in Electrical Engineering from Osmania University, Hyderabad, India, in 1973; M.Tech. degree in Power systems from Jawaharlal Nehru Technological University, Hyderabad, India, in 1976; and Ph.D. degree in Control Systems from the Indian Institute of Technology (IIT) Delhi, New Delhi, India, in 1981. From April to November 1988, he was with the Machine Vision Group, City University London, London, UK, and from March to June 1993, he was with the Robotics Research Group, Oxford University, Oxford, UK. His current research interests include fuzzy modeling for dynamic systems and applications of fuzzy logic to image processing, medical imaging, multimodal biometrics and surveillance. He has authored a book on Computer Graphics in 2005 under PBP publications and also has well over 325 publications in both conferences and journals to his credit. He has guided 32 Ph.Ds. and 160 M.Tech students. He has handled several sponsored projects. He was an Associate Editor of both Pattern Recognition Journal (January 2005–March 2011) and of IEEE Transactions on Fuzzy Systems and a viewer to other journals such as Pattern Recognition Letters, IEEE Transactions on Image Processing and Systems, Man and Cybernetics (January 2007–2010). He is a Senior member of IEEE and is listed in Reference Asia; Asia's who's who of Men and Women of achievement; 5000 Personalities of the World (1998), American Biographical Institute. He was a Guest Editor of Defense Science Journal for the special issue on "Information sets and Information Processing" September, 2011.

Condenser Heat Recovery For Combined Cooling-Heating and Power Generation using Isentropic Fluid

Kaushalendra K. Dubey, and R. S. Mishra

***Abstract*—The thermodynamic study of combined cooling, heating and power generation system with the recovery of heat source from the condenser of 1MW thermal power plant. The R-134a working fluid based Organic Rankine Cycle (ORC) is introduced in the proposed thermodynamic analysis and the provision of parabolic trough collector is recommended for solar heating purpose. The analysis of the system shows thermal efficiency and multiple effects like Heating- Cooling and Power through heat recovery of thermal power plant. This analysis also express the resulting process heat obtained and the cooling effect of solar integrated as well as non-solar integrated system. The results conclude that the thermal efficiency as well as heating-power and cooling effects increases by 30-40% in the case of solar integrated as compare to without solar system.**

***Index Terms*—Organic Rankine Cycle; Solar Energy; Process Heat; Cooling Effect; Expander work.**

I. Introduction

EIA observed that the increasing rate of consumption of fossil fuels to fulfill industrial, commercial and domestic power requirement. More than 38% (6629 trillion BTU) and 8% (5149 trillion BTU) of the power generated in the world is consumed by the petroleum and chemical industry in US. The negative impact on the environment is in the form of higher amount of pollutants & toxic emissions, which are released into the atmosphere. It is estimated that over the next century this high rate of consumption will be majorly responsible for the emission of CO_2 and other gases that cause global warming through greenhouse effect. It is, therefore, now the responsibility of every country in the world to improve the quality of their sources of energy in order to reduce CO_2 emission and oil dependency [1]. The utilization of dumped heat (waste heat) into atmosphere from power plants, industry like glass, cement, textile, etc by method of heat recovery coupled with Organic Rankine Cycle (ORC) based renewable energy conversion system for combined effect of cooling-heating and power. The ORC applies the principle of the steam Rankine cycle but uses an organic working fluid, with low boiling point, instead of steam to recover heat from a lower temperature source. ORC is a capable technology for the conversion of these low temperature sources to power [2]-[5]. The energy recovery from low grade heat sources for work or for production of electricity is possible in ORC because of the use of an organic fluid. Thus, one of the main factors of research identified with relation to this issue is the selection of a suitable working fluid since it has a great influence on the thermodynamic analysis and process [6]-[10]. The selection of organic working fluid is depending on the application, the heat source and the range of temperature for the application, working fluid has optimum thermodynamic properties at the lowest possible temperatures and pressures. It must also satisfy several other criteria such as being economic, non-toxic, non- flammable and environmentally friendly, thus, allowing high utilization of the energy available from the heat sources. Riccardo Vesocove survey demonstrated waste heat recovery by using ORC in cement, glass and steel industry as shown in Table I.

II. R134A-Isentropic Fluid for Heat Recovery Application

For heat recovery applications, a group of isentropic type fluid R 134a is applied in ORC system. In ORC, superheating is not necessary while applying isentropic fluids because its expansion would not cause any wet fluid problem. Some fluids likeR-141b, R-123, R-21, R-245ca, R-245fa, R-236ea and R-142b etc, have critical temperature above 400 K, resulting to be used only in supercritical cycle for low grade heat sources [9]. R134a is belongs from Group-A1 of ASHRAE, that means nontoxic and nonflammable fluid with zero Ozone Depletion Potential (ODP) value. R134a have wide variety of turbo-machines, heat exchangers. Thermo-physical properties of R134a has a high molecular mass (chemical expression; CF3CH2F), 102kg/kmol. It has a temperature and critical pressure of 101.1ºC and 40.6 bar, respectively, and allowing it to operates at a higher pressure than atmospheric pressure at condenser part, which prevents any air leakage. Several articles have reported the application and implementation of ORC with R134a as the working fluid.

TABLE I: ORC Waste Heat Recovery in Industries[1]

Industry/Application	Cement	Glass	Steel (Roll milling)	Unit
Heat Sources	Kiln and Clinker	Oven exhaust gas	Preheating exhaust	
Plant Capacity	2500	500	6000	Tons / day
Waste thermal power in exhaust gas	12	5	13	MW
ORC thermal power	11	4.7	0	MW
Net electric generation through ORC	1.6	1	2.4	MW
ORC Cost	1.8	1.3	2.4	Million €
Avoided CO2 emission	9664	5520	12096	Tons / year

Published on June 16, 2017.

K. K. Dubey is associated with the Dept of Mechanical Engg, School of Engg & Technology,Sharda University,GreaterNoida,State-UP,India, and research fellow of Delhi Technological University, Delhi, India (e-mail: dubey.kaushalendra@gmail.com).

Dr R.S. Mishra is with Dept of Mechanical Engineering, Delhi Technological University, Delhi, India (e-mail: professor_rsmishra@yahoo.co.in)

[1] ORC recovering industrial heat- Power generation from waste heat energy stream report by Riccardo Vascovo, Turboden, Brecia, Italy, email-riccardo.vescovo@turboden.it

The other fluids such as benzene, ammonia, R134a, R113, R11 and R12 was studied for ORC [8],[9],[11]. The parametric optimization of a regenerative ORC using R12, R123, R134a, and R717 as working fluids exposed that R123 as working fluid performs to be a good system for converting low-grade waste heat to power. A low temperature (35-70 ^{0}C) solar integrated Organic Rankine Cycle system built with R134a for reverse osmosis desalination in Greece. The similar results of thermal efficiencies 4-7% was achieved as previous analysis of desalination systems [12]-[23]. Simulation based study of two stage ORC available for desalination of an upper ORC coupled with a lower ORC with R134a as the working fluid obtained an efficiency of 4.2% and a global efficiency of 3%. Furthermore, other cycles with R134a as the working fluid for applications in geothermal sources and in bottoming cycles with internal combustion engine have been reported [24]-[27]. A thermodynamic inspection of 31 pure component working fluids for organic Rankine cycle has shown an efficiency of 7.7% being achieved in cycles that operate with R134a at temperature of 100.0ºC and another fluids were studied and it was reported that R134a is the most suitable [4],[6]. The thermodynamic modeling of ORC based micro CHP discussed efficiency, water flow rate and achievable electricity production rate [28]. Thermodynamic evaluation of basic ORC with arrangement of internal heat exchanger examined by Fredy Vélez et al estimated to be 11-14% of maximum efficiency at conditions of saturation and overheating of working fluid, and also conclude the effect of pressure ratio and overheating of fluid in performance of system [33]. Other works that have analyzed the use of R134a as working fluid in ORC for low grade heat recovery. There is a great interest in the use of this fluid for generation of energy in systems with heat input temperature below 150ºC [29]-[32]. However, we noticed in the literature about the best thermodynamic conditions for its use and a gradual impact has been observed working with R134a fluid at superheated conditions. In the influence of the source of energy on the performance of the system as well as on the net specific work output derived from the system. The T-S and T-h property of R-134a is shown in Fig. 1 and 2.

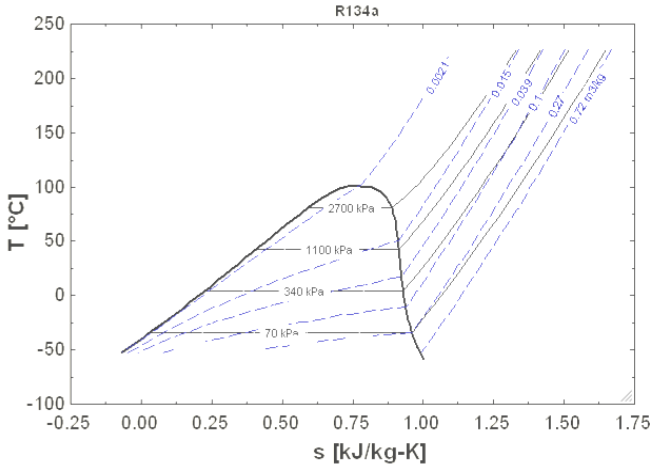


Fig. 1. T-S plot of R-134a- EES

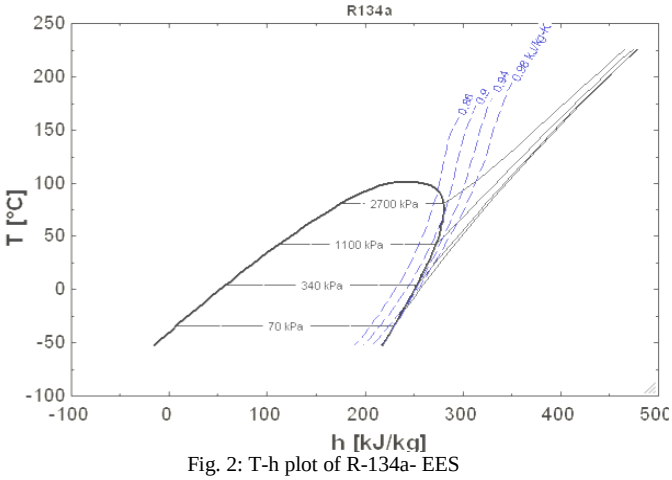


Fig. 2: T-h plot of R-134a- EES

III. Energy Recovery (High-Low Grade) Technology

Three categories of waste heat sources are identified with respect to the temperature range. Sets of waste heat recovery and their corresponding thermal ranges are listed in Table II.

TABLE II: Grade of Energy and its Heat Recovery Technology [3]

Heat Categories	Heat Sources	Temperature in ^{0}C	Suggested recovery technology
High Grade (>650^{0}C)	Solid waste	650-1000	Air preheating
	Nickel refining furnace	1370-1650	Steam generation for heating
	Copper reverberatory furnace	760-815	Thermoelectric and thermal PV
	Glass melting furnace	1000-1550	Heat exchanger for preheating
	Hydrogen plant	650-1000	Thermal PV
Medium Grade (230-650^{0}C)	Steam boiler exhaust	230-480	Steam Rankine cycle
	Gas turbine exhaust	370-540	Organic Rankine Cycle
	Drying and baking ovens	230-600	Thermal PV
	Catalytic crackers	425-650	Thermal PV
	Reciprocating engine exhaust	315-600	Thermoelectric
Low Grade (>230^{0}C)	Welding and injection molding	32-88	Kalina cycle
	Hot processed liquids and solids	32-233	Organic Rankine cycle
	Drying, Baking and Curing ovens	93-230	Absorption and adsorption cooling
	Bearing	32-88	Piezoelectric

IV. Cycle for Combined Cooling Heating and Power Generation by Heat Recovery- Working Principle and Schematic Description

Fig. 3 clearly explains all the systems included in the proposed model. It consists of a refrigerant storage from which the refrigerant is introduced into the system in form of superheated vapour. Any makeup of refrigerant, if required, during the working of the system is also obtained from the refrigerant tank. The outlet of the refrigeration tank is provided with a two-way flow control valve to regulate the mass flow rate of the superheated refrigerant vapour into the Organic Rankine Cycle. The superheated refrigerant vapour then enters the condenser heat exchanger, which enables condenser heat to be recovered. Here the condenser heat is transferred to the refrigerant and at the ORC outlet of the condenser heat exchanger, superheated refrigerant vapour, in supercritical condition, is obtained which then passes through expander 1. Expander 1 is coupled to a generator, which as a result of the work done in expander 1

produces electric power. The superheated refrigerant vapour obtained from the outlet of expander 1 is at a lower temperature and enters into HE 1 as the hot fluid. Air having ambient temperature is used as the cold fluid in HE 1. A two-way flow control valve is provided at the air inlet of HE 1 to regulate the mass flow rate of air into it. Heat, which is rejected from the superheated refrigerant vapour, is added to the air in HE 1. The hot air obtained at the outlet of HE 1 is sent to fulfill industrial process heat requirement. A tapping is provided from the air outlet of HE 1 for extracting hot air to reheat the superheated refrigerant vapour later in HE 3, if required. A three-way flow control valve is provided at the tapping junction of the air outlet of HE 1 to regulate the mass flow rate of the hot air being extracted, for reheating, from that going for industrial heating process. The superheated refrigerant vapour obtained from the ORC outlet of HE 1 then passes through the throttle where its pressure and temperature are drastically dropped. The superheated refrigerant vapour then passes through HE 2. Air having ambient temperature is used as the hot fluid in HE 2. A two-way flow control valve is provided at the air inlet of heat HE 2 to regulate the mass flow rate of air into it. Heat, which is rejected from the air, is added to the superheated refrigerant vapour in the heat exchanger. The cold air obtained at the air outlet of HE 1 is sent for space cooling or chilling. The superheated refrigerant vapour obtained at the ORC outlet of HE 2, at a higher temperature, passes through HE 3. The mixture of hot air extracted from the outlet of HE 1 and HE 4 is used as the hot fluid in HE 3. A three-way flow control valve is provided at the junction of the two extraction lines to regulate the mass flow rate of the hot air from HE 1 and HE 4 going into the mixture. Heat is rejected from the hot air and added to the superheated refrigerant vapour in the heat exchanger. The cooled air is released into the atmosphere from the outlet of HE 3. A bypass line is provided across the solar heat exchanger to bypass it in case of the condition of the solar heater and the solar heat exchanger not being in use. Another bypass line is provided bypassing expander 2 and HE 4 in case of the same. A three-way flow control valve is provided at the inlet of the solar heat exchanger bypass line to enable or disable solar heat exchanger bypass. Another three-way flow control valve is provided at the outlet of the solar heat exchanger bypass line to prevent the back flow of the superheated refrigerant vapour, coming out of the solar heat exchanger, into the bypass line of solar heat exchanger, in case the solar heat exchanger is in use, or to prevent the flow of the superheated refrigerant vapour into the outlet line of the solar heat exchanger, in case the solar heat exchanger is not in use. A three-way flow control valve is provided at the inlet of the expander 2 - HE 4 bypass line to bypass these systems when solar heat exchanger is not in use. Another three-way flow control valve is provided at the outlet of the expander 2- HE 4 bypass line to prevent the flow of the superheated refrigerant vapour, coming out of HE 4, into the expander 2 - HE 4 bypass line, in case the solar heat exchanger is in use, or to prevent the back flow of the superheated refrigerant vapour into the outlet line of HE 4, in case the solar heat exchanger is not in use. The designed model operates with solar and without solar thermal integration. Schematic of the proposed model for combined cooling, heating and power generation by condenser heat recovery has been made using the software AUTOCAD 2016 and all points have described in Table III.

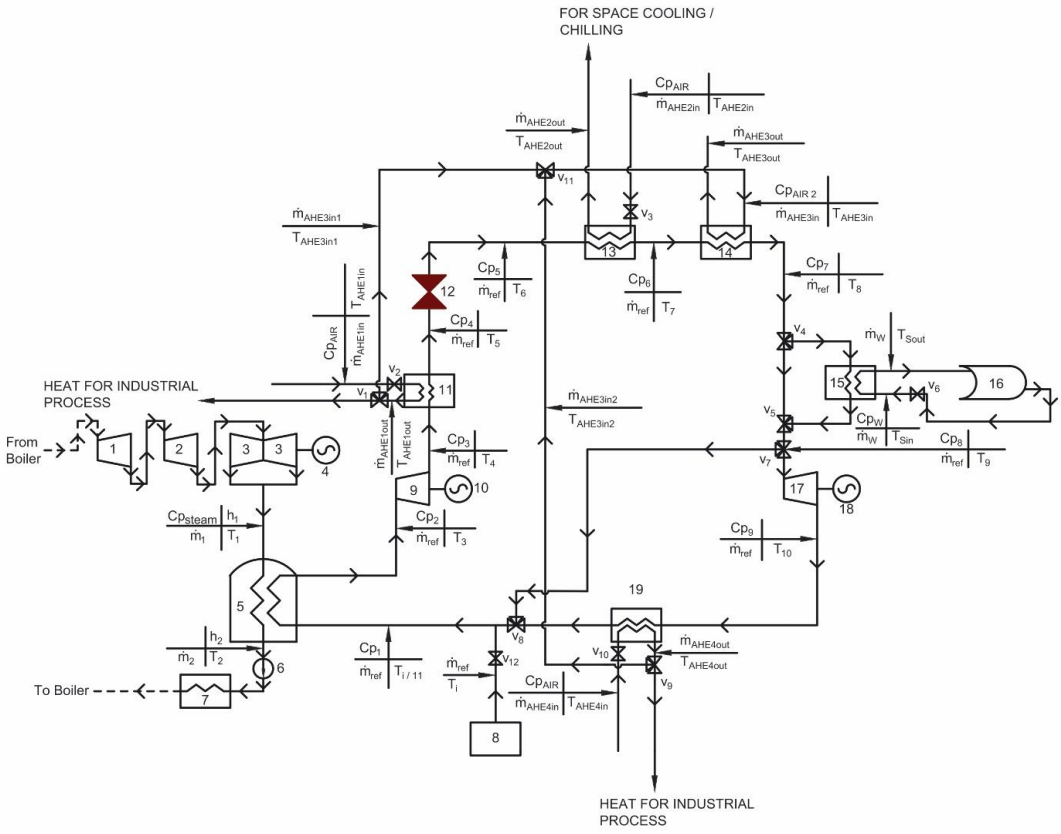


Fig. 3. Schematic of Combined Cooling, Heating and Power Generation Plant using condenser heat recovery

TABLE III: DESCRIPTION OF SYSTEMS USED IN THE SCHEMATIC (FIG. 3)

System number	System description	System number	System description	System number	System description
1	High pressure turbine	12	Throttle	V_4	Three way flow control valve
2	Intermediate pressure turbine	13	Heat exchanger 2	V_5	Three way flow control valve
3	Low pressure turbine	14	Heat exchanger 3	V_6	Two way flow control valve
4	Power generator	15	Solar heat exchanger	V_7	Three way flow control valve
5	Condenser	16	Solar heater	V_8	Three way flow control valve
6	Circulation Extraction pump	17	Expander 2	V_9	Three way flow control valve
7	Feedwater Heater	18	Power generator	V_{10}	Two way flow control valve
8	Refrigerant tank	19	Heat exchanger 4	V_{11}	Three way flow control valve
9	Expander 1	V_1	Three way flow control valve	V_{12}	Two way flow control valve
10	Power generator	V_2	Two way flow control valve		
11	Heat exchanger 1	V_3	Two way flow control valve		

V. THERMODYNAMIC ANALYSIS OF PROPOSED MODEL

The steam from the outlet of low pressure turbine of the Rankine cycle enters the condenser at temperature T_1, having mass flow rate $\dot{m}_1$, enthalpy h_1 and temperature T_1. Condensed water leaves the condenser at temperature T_2 having mass flow rate $\dot{m}_2$ (= $\dot{m}_1$) and enthalpy h_2. Knowing the values of the above parameters, and taking the specific heat capacity of the saturated steam as Cp_{steam}, the heat rejection by the steam in the condenser heat exchanger is calculated using heat balance equation

$$Q_{COND} = \dot{m}_1 Cp_{steam} (T_1 - T_2) \quad (1)$$

Taking the latent heat rejected by the steam in the condenser heat exchanger as Q_{LH} and the condenser heat loss 15%, the total heat available to the refrigerant vapour is calculated from:

$$Q_{Total} = \{Q_{COND} + (Q_{LH} * \dot{m}_1)\} - [\{Q_{COND} + (Q_{LH} * \dot{m}_1)\} * 0.15] \quad (2)$$

Superheated refrigerant (R 134a) vapour from the refrigerant storage is introduced into the system at a known temperature T_i for the startup cycle. In all operation cycle the refrigerant vapour enters the condenser heat exchanger at temperature T_{11}. Knowing the value of the specific heat capacity, Cp_1, of the superheated refrigerant vapour at the inlet of the condenser heat exchanger and taking different values for mass flow rate, $\dot{m}_{ref}$, of the same, ORC outlet temperature, T_3, of the superheated refrigerant vapour is calculated using heat balance equation:

$$Q_{Total} = \dot{m}_{ref} Cp_1 (T_1 - T_{i/11}) \quad (3)$$

The mass flow rate, $\dot{m}_{ref}$, of the superheated refrigerant vapour is optimised from the data generated.

Superheated refrigerant vapour leaves the condenser heat exchanger and enters expander 1 at temperature T_3, having mass flow rate $\dot{m}_{ref}$and specific heat capacity Cp_2. It leaves expander 1 at temperature T_4, having mass flow rate $\dot{m}_{ref}$and specific heat capacity Cp_3. Knowing the value of the specific heat capacity, Cp_2 and taking different values for efficiency of expander 1 (η_{E1}), T_4 and the work done by expander 1 (W_{E1}) is calculated using the equations:

$$\eta_{E1} = W_{E1} / Q_{Total} \quad (4)$$

Where;

$$W_{E1} = \dot{m}_{ref} Cp_2 (T_3 - T_4) \quad (5)$$

Superheated refrigerant vapour leaves expander 1 and enters heat exchanger 1 as the hot fluid at temperature T_4 and having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_3. It leaves heat exchanger 1 at temperature T_5, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_4. Air enters heat exchanger 1 as the cold fluid at temperature T_{AHE1in}, having mass flow rate $\dot{m}_{AHE1in}$ and specific heat capacity Cp_{AIR}. It leaves heat exchanger 1 at temperature $T_{AHE1out}$ and having mass flow rate $\dot{m}_{AHE1out}$ (= $\dot{m}_{AHE1in}$). The values of T4, $\dot{m}_{ref}$, Cp_3 and Cp_{AIR} are known. T_{AHE1in} is taken as 288.15 K as per standard atmosphere data. Taking the heat exchanger heat loss as 15%, T_5, $T_{AHE1out}$ and heat available in heat exchanger 1 (Q_{HE1}) for different values of mass flow rate, $\dot{m}_{AHE1in}$, are calculated using heat balance equation:

$$Q_{HE1} = \dot{m}_{ref} Cp_3 (T_4 - T_5) - [\{\dot{m}_{ref} Cp_3 (T_4 - T_5)\} * 0.15] \quad (6)$$

$$Q_{HE1} = \dot{m}_{AHE1in} Cp_{AIR} (T_{AHE1out} - T_{AHE1in}) \quad (7)$$

The mass flow rate $\dot{m}_{AHE1in}$ of the air entering heat exchanger 1 is optimised from the data generated.

Superheated refrigerant vapour leaves heat exchanger 1 and enters the throttle at temperature T_5 and having mass flow rate $\dot{m}_{ref}$. It leaves the throttle at temperature T_6 and having mass flow rate $\dot{m}_{ref}$. The value of T_6 is optimised based on the properties of the throttle.

Superheated refrigerant vapour leaves the throttle and enters heat exchanger 2 as the cold fluid at temperature T_6, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_5. It leaves heat exchanger 2 at temperature T_7, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_6. Air enters heat exchanger 2 as the hot fluid at temperature T_{AHE2in}, having mass flow rate $\dot{m}_{AHE2in}$ and specific heat capacity Cp_{AIR}. It leaves heat exchanger 2 at temperature $T_{AHE2out}$ and having mass flow rate $\dot{m}_{AHE2out}$ (=$\dot{m}_{AHE2in}$). The values of T_6, $\dot{m}_{ref}$, Cp_5 and Cp_{AIR} are known. T_{AHE2in} is taken as 288.15 K as per standard atmosphere data. Taking the heat exchanger heat loss as 15%, T_7, $T_{AHE2out}$ and heat available in heat exchanger 2 (Q_{HE2}) for different values of mass flow rate of air ($\dot{m}_{AHE2in}$) are calculated using heat balance equation:

$$Q_{HE2} = \dot{m}_{ref} Cp_5 (T_7 - T_6) - [\{\dot{m}_{ref} Cp_5 (T_7 - T_6)\} * 0.15] \quad (8)$$

$$Q_{HE2} = \dot{m}_{AHE2in} Cp_{AIR} (T_{AHE2in} - T_{AHE2out}) \quad (9)$$

The mass flow rate $\dot{m}_{AHE2in}$ of the air entering heat exchanger 2 is optimised from the data generated.

Superheated refrigerant vapour leaves heat exchanger 2 and enters heat exchanger 3 as the cold fluid at temperature T_7, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_6. It leaves heat exchanger 3 at temperature T_8, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_7. Air enters heat exchanger 3 as the hot fluid at temperature T_{AHE3in}, having mass flow rate $\dot{m}_{AHE3in}$ and specific heat capacity Cp_{AIR2}. It leaves heat exchanger 3 at temperature $T_{AHE3out}$ and having mass flow rate $\dot{m}_{AHE3out}$ (= $\dot{m}_{AHE3in}$). $\dot{m}_{AHE3in}$ is calculated using heat balance equation:

$$\dot{m}_{AHE3in} = \dot{m}_{AHE3in1} + \dot{m}_{AHE3in2} \quad (10)$$

The values of T_7, $\dot{m}_{ref}$, Cp_6 and Cp_{AIR2} are known. T_{AHE3in} is calculated using the equation

$$\dot{m}_{AHE3in1}\, Cp_{AIR}\, (T_{AHE3in1} - T_{AHE3in}) = \dot{m}_{AHE3in2}\, Cp_{AIR}\, (T_{AHE3in} - T_{AHE3in2}) \quad (11)$$

Taking the heat exchanger heat loss as 15%, T_8, $T_{AHE3out}$ and heat available in heat exchanger 3 (Q_{HE3}) for different values of mass flow rate of air ($\dot{m}_{AHE3in}$) are calculated using the equation:

$$Q_{HE2} = \dot{m}_{AHE3in}\, Cp_{AIR2}\, (T_{AHE3in} - T_{AHE3out}) - [\{\dot{m}_{AHE3in}\, Cp_{AIR2}\, (T_{AHE3in} - T_{AHE3out}) * 0.15] \quad (12)$$

$$Q_{HE2} = \dot{m}_{ref}\, Cp_6\, (T_8 - T_7) \quad (13)$$

$\dot{m}_{AHE3in1}$and $\dot{m}_{AHE3in2}$ are optimised from the data generated.

A. *Case I: With solar operation*

Superheated refrigerant vapour leaves heat exchanger 3 and enters solar heat exchanger as the cold fluid at temperature T_8, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_7. It leaves solar heat exchanger at temperature T_9, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_8.Water enters solar heat exchanger as the hot fluid at temperature T_{Sin}, having mass flow rate $\dot{m}_W$ and specific heat capacity Cp_W. It leaves the solar heat exchanger at temperature T_{Sout} and having mass flow rate $\dot{m}_W$. The values of T_8, $\dot{m}_{ref}$, Cp_7 and Cp_W are known. Taking the heat exchanger heat loss as 15%, T_9, T_{Sout} and heat available in solar heat exchanger (Q_{Solar}) for different values of mass flow rate of water ($\dot{m}_W$) is calculated using heat balance equation:

$$Q_{Solar} = \dot{m}_W Cp_W\, (T_{Sin} - TSout) - [\{\dot{m}_W Cp_W\, (T_{Sin} - T_{Sout})\} * 0.15] \quad (14)$$

$$Q_{Solar} = \dot{m}_{ref}\, Cp_7\, (T_9 - T_8) \quad (15)$$

The mass flow rate $\dot{m}_W$ of the water entering the solar heat exchanger is optimised from the data generated.

Superheated refrigerant vapour leaves the condenser heat exchanger and enters expander 2 at temperature T_9, having mass flow rate $\dot{m}_{ref}$and specific heat capacity Cp_8. It leaves expander 2 at temperature $T_{10,}$ having mass flow rate $\dot{m}_{ref}$and specific heat capacity Cp_9. Knowing the value of the specific heat capacity, $Cp_{8,}$ of the superheated refrigerant vapour at the inlet of expander 1 and taking different values for efficiency of expander 1(η_{E1}) T_{10} and the work done by expander 2 (W_{E2}) is calculated using the equations:

$$\eta_{E2} = W_{E2} / Q_{Solar} \quad (16)$$

Where;

$$W_{E2} = \dot{m}_{ref}\, Cp_8\, (T_9 - T_{10}) \quad (17)$$

Superheated refrigerant vapour leaves expander 2 and enters heat exchanger 4 as the hot fluid at temperature T_{10} and having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_9. It leaves heat exchanger 4 at temperature T_{11}, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_1.The value of Cp_1 will change in the transition from startup cycle to operation cycle. Air enters heat exchanger 2 as the cold fluid at temperature T_{AHE4in}, having mass flow rate $\dot{m}_{AHE4in}$ and specific heat capacity Cp_{AIR}. It leaves heat exchanger 4 at temperature $T_{AHE4out}$ and having mass flow rate $\dot{m}_{AHE4out}$ (= $\dot{m}_{AHE4in}$). The values of T_{10}, $\dot{m}_{ref}$, Cp_9 and Cp_{AIR} are known. T_{AHE4in} is taken as 288.15 K as per standard atmosphere data. Taking the heat exchanger heat loss as 15%, assuming that no makeup is required, T_{11}, $T_{AHE4out}$ and heat available in heat exchanger 4 (Q_{HE4}) for different values of mass flow rate of air ($\dot{m}_{AHE4in}$) are calculated using heat balance equation:

$$Q_{HE4} = \dot{m}_{ref}\, Cp_9\, (T_{11} - T_{10}) - [\{\dot{m}_{ref}\, Cp_9\, (T_{11} - T_{10})\} * 0.15] \quad (18)$$

$$Q_{HE4} = \dot{m}_{AHE4in}\, Cp_{AIR}\, (T_{AHE4out} - T_{AHE4in}) \quad (19)$$

The mass flow rate $\dot{m}_{AHE4in}$ of the air entering heat exchanger 4 is optimised from the data generated. The entire cycle then repeats as long as the system is operating.

B. *Case II: Without solar operation:*

Superheated refrigerant vapour leaves heat exchanger 3 and enters condenser heat exchanger as the cold fluid. Assuming that no makeup is required, it enters the condenser heat exchanger at temperature T_{11} that is equal to T_8, having mass flow rate $\dot{m}_{ref}$ and specific heat capacity Cp_1 that is equal to Cp_7. The value of Cp_1 will change in the transition from startup cycle to operation cycle. The entire cycle then repeats as long as the system is operating.

VI. Results and Discussion

Increasing the mass flow rate of air in HE2 increases the heat obtained thereby reducing the cooling effect and heating the air more. In operation cycle with solar heater the change in heat obtained and the rise in air temperature for a small increase in mass flow rate of air is much more than that in the operation cycle without solar heater. Fig 4 shows integrating solar heating with operation cycle increases the process heat obtained for industrial purpose.

Fig 5 depicts the efficiency of scroll type expander is in the range of 14% to 25%. The expander in case of cycle without solar integration will do more work. In case of solar integration, more heat for industrial process will be obtained.

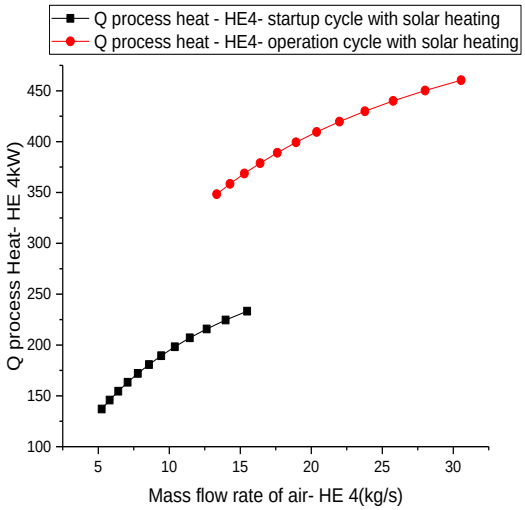


Fig. 4. Effect of mass flow rate of air on process heat- HE 4

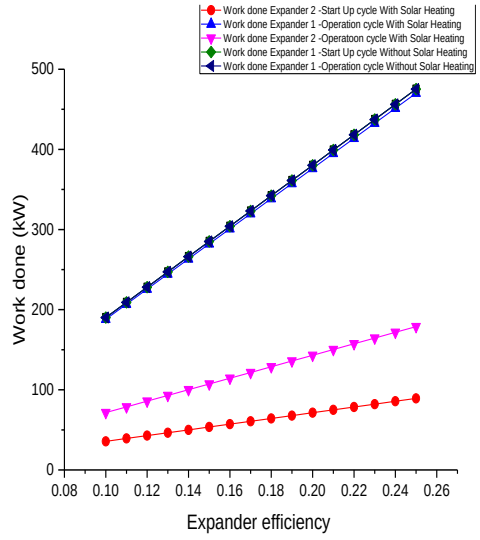


Fig. 5. Effect of expander efficiency on work done by expander

Fig 6 explains thatwhen solar heating is not implemented, the heat for process obtained from HE 1 will be more in both cycles than that obtained when solar heating is implemented. Moreover, the gain in the heat obtained from HE 1, for process, per unit increase of mass flow rate of air passing through the HE 1, in case of operation cycle, is greater when solar heating is not implemented than when solar heating is implemented. All parametric results are concluded in Table IV.

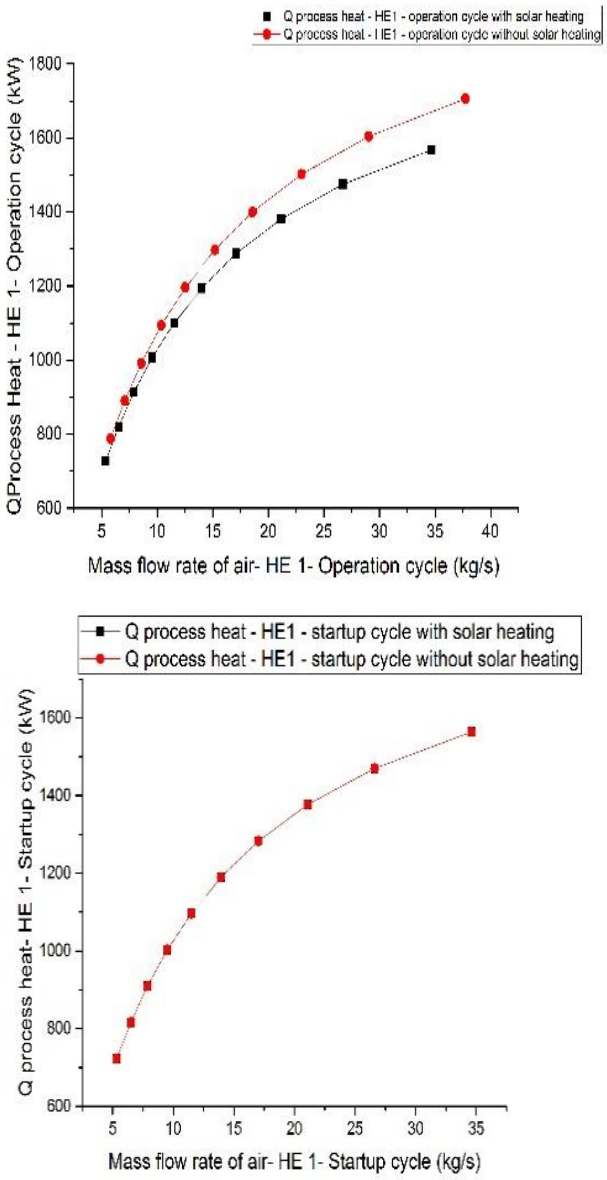


Fig.6. Effect of mass flow rate of air, flowing through HE 1, on Q process heat in operation cycle and startup cycle

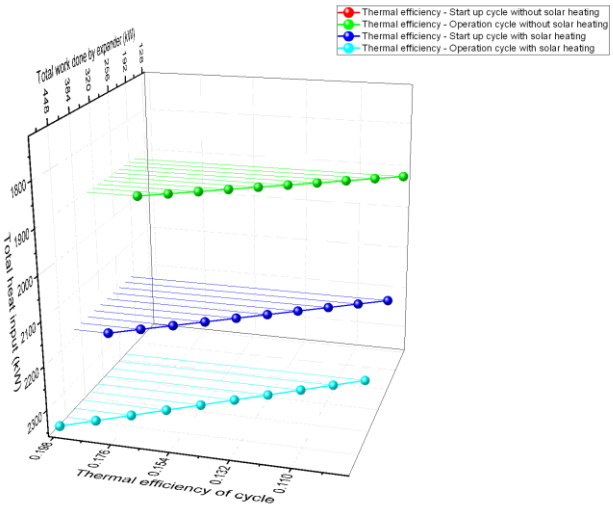


Fig. 7. Total work done by expander Vs. Total heat input Vs. Thermal efficiency of cycle

TABLE IV: PARAMETRIC RESULTS

Parametric Results	Cycle without solar	Cycle with solar
$\eta_{thermal}$	8-15%	12-27%
$\eta_{expander}$	10-25%	10-25%
Process Heat ($Q_{process\ heat}$)	900-1100 kW	1200-1500 kW
Cooling Effect ($Q_{Cooling}$)	14- 19.25 kW	35- 38.5 kW
Expander work ($W_{expander}$)	200-470 kW	260-650 kW
Solar Heating (Q_{Solar})	N/A	460 kW
Solar Collector Efficiency	N/A	20-25%
Water Temp through solar field	N/A	343.15-363.15 K

Fig. 7 shows the expander work obtained and the thermal efficiency will be more in case of startup and operation cycle without solar heating. This is because in case of startup and operation cycle without solar heating, work output is being provided by only one expander and also heat obtained for industrial process is only from one heat exchanger as opposed to the case of startup and operation cycle with solar heating where the work output is being provided by two expanders and the heat obtained for industrial process is only from two heat exchanges with additional heat input to the cycle through the utilization of solar energy.

VII. CONCLUSION

Increasing the mass flow rate of air increases the rate of heat exchange, thereby raising the temperature of the cold refrigerant more rapidly per unit time. This increase the amount of heat rejected by air or Qcooling but has a negative effect on the decrement of the air temperature leading to an increase in the temperature of air at the outlet of HE 2 due to upward shift of the equilibrium point of heat exchange. Important conclusions of proposed analysis mentioned below and parametric results shown in Table IV.

1. For effective cooling of 278.15 K or below, the mass flow rate of air through HE-2 should not exceed 8-10 kg/s.
2. Cycle with solar coupling provides more heat for process. This is nearly two times that of the heat provided for process by the cycle without solar coupling, but the total work achieved in case of solar heating is less than that achieved without solar heating.
3. More power from the expanders is achieved when solar heating is not implemented. This is due to the reason that the heat generated by the solar heater is converted into process heat. Therefore, solar integration is recommended, for the purpose of obtaining more process heat, in summer and for the purpose of power generation in winter.
4. In cooling effect achieved in the operation cycle in both cases, that is, with solar integration and without solar integration, is the same. The same stands true also for the cooling effect achieved in the startup cycle in both cases.
5. The refrigerants specific heat capacity will vary with temperature and pressure.
6. The parametric table concludes that the performance indicators (thermal efficiency, SCP, heating and cooling effect) of proposed system is increased by 30-40% using solar system.

APPENDIX

η_{E1}= Efficiency of expander 1
η_{E2}= Efficiency of expander 2
$\eta_{thermal}$= Thermal efficiency of the Organic Rankine Cycle
$\eta_{expander}$= Combined efficiency of the expanders
CO_2= Carbon Dioxide
Cp_{steam}=Specific heat capacity of steam at the inlet of condenser [kJ/kg-K]
Cp_1=Specific heat capacity of refrigerant at the inlet of condenser heat exchanger [kJ/kg-K]
Cp_2=Specific heat capacity of refrigerant at the inlet of expander 1 [kJ/kg-K]
Cp_3=Specific heat capacity of refrigerant at the inlet of heat exchanger 1 [kJ/kg-K]
Cp_4=Specific heat capacity of refrigerant at the inlet of throttle [kJ/kg-K]
Cp_5=Specific heat capacity of refrigerant at the inlet of heat exchanger 2 [kJ/kg-K]
Cp_6= Specific heat capacity of refrigerant at the inlet of heat exchanger 3 [kJ/kg-K]
Cp_7= Specific heat capacity of refrigerant at the outlet of heat exchanger 3 [kJ/kg-K]
Cp_8= Specific heat capacity of refrigerant at the inlet of expander 2 [kJ/kg-K]
Cp_9= Specific heat capacity of refrigerant at the inlet of heat exchanger 4 [kJ/kg-K]
Cp_{AIR}= Specific heat capacity of air at the inlet of heat exchanger 1, 2,3, 4 [kJ/kg-K]
$Cp_{AIR\ 2}$= Specific heat capacity of air at the inlet of heat exchanger 3 [kJ/kg-K]
Cp_w= Specific heat capacity of refrigerant at the inlet of solar heat exchanger [kJ/kg-K]
h_1= Enthalpy of saturated steam at the inlet of condenser [kJ/kg]
h_1= Enthalpy of water at the outlet of condenser [kJ/kg]
HE 1= Heat exchanger 1
HE 2= Heat exchanger 2
HE 3= Heat exchanger 3
HE 4= Heat exchanger 4
$\dot{m}_1$= Mass flow rate of saturated steam at condenser inlet [kg/s]
$\dot{m}_2$= Mass flow rate of water at condenser outlet [kg/s]
$\dot{m}_{ref}$= Mass flow rate of refrigerant [kg/s]
$\dot{m}_{AHE1in}$= Mass flow rate of air at the inlet of heat exchanger 1 [kg/s]
$\dot{m}_{AHE1out}$= Mass flow rate of air at the outlet of heat exchanger 1 [kg/s]
$\dot{m}_{AHE2in}$= Mass flow rate of air at the inlet of heat exchanger 2 [kg/s]
$\dot{m}_{AHE2out}$= Mass flow rate of air at the outlet of heat exchanger 2 [kg/s]
$\dot{m}_{AHE3in1}$= Mass flow rate of extraction taken from outlet of heat exchanger 1 [kg/s]
$\dot{m}_{AHE3in2}$= Mass flow rate of extraction taken from outlet of heat exchanger 4 [kg/s]
$\dot{m}_{AHE3in}$= Mass flow rate of air at the inlet of heat exchanger 3 [kg/s]
$\dot{m}_{AHE3out}$= Mass flow rate of air at the outlet of heat exchanger 3 [kg/s]
$\dot{m}_{AHE4in}$= Mass flow rate of air at the inlet of heat exchanger 4 [kg/s]
$\dot{m}_{AHE4out}$= Mass flow rate of air at the outlet of heat exchanger 4 [kg/s]

$\dot{m}_w$= Mass flow rate of water at the inlet and outlet of solar heat exchanger [kg/s]
ORC= Organic Rankine Cycle
PV= Photovoltaic
Q_{COND}= Heat rejected by the steam in the condenser heat exchanger without considering the heat exchanger losses [kW]
Q_{Total}= Total heat available to the refrigerant vapour in condenser heat exchanger after considering the heat exchanger losses [kW]
Q_{HE1}= Heat available for exchange in heat exchanger 1 [kW]
Q_{HE2}= Heat available for exchange in heat exchanger 2 [kW]
Q_{HE3}= Heat available for exchange in heat exchanger 3 [kW]
Q_{HE4}= Heat available for exchange in heat exchanger 4 [kW]
Q_{Solar}= Heat available for exchange in solar heat exchanger [kW]
$Q_{Cooling}$= Heat rejected by air in heat exchanger 2 [kW]
$Q_{Process\ heat}$= Total heat available for industrial processes [kW]
T_i= Temperature of refrigerant at the inlet of condenser heat exchanger for startup cycle [K]
T_1= Temperature of saturated steam at the inlet of condenser [K]
T_2= Temperature of water at the outlet of condenser [K]
T_3= Temperature of refrigerant at the inlet of expander 1 [K]
T_4= Temperature of refrigerant at the inlet of heat exchanger 1 [K]
T_5= Temperature of refrigerant at the inlet of throttle [K]
T_6= Temperature of refrigerant at the inlet of heat exchanger 2 [K]
T_7= Temperature of refrigerant at the inlet of heat exchanger 3 [K]
T_8= Temperature of refrigerant at the outlet of heat exchanger 3 [K]
T_9= Temperature of refrigerant at the inlet of expander 2 [K]
T_{10}= Temperature of refrigerant at the inlet of heat exchanger 4 [K]
T_{11}= Temperature of refrigerant at the inlet of condenser heat exchanger for operation cycle [K]
T_{AHE1in}= Temperature of air at the inlet of heat exchanger 1 [K]
$T_{AHE1out}$= Temperature of air at the outlet of heat exchanger 1 [K]
T_{AHE2in}= Temperature of air at the inlet of heat exchanger 2 [K]
$T_{AHE2out}$= Temperature of air at the outlet of heat exchanger 2 [K]
$T_{AHE3in1}$= Temperature of the extraction taken from the outlet of heat exchanger 1 [K]
$T_{AHE3in2}$= Temperature of the extraction taken from the outlet of heat exchanger 4 [K]
T_{AHE3in}= Temperature of air at the inlet of heat exchanger 3 [K]
$T_{AHE3out}$= Temperature of air at the outlet of heat exchanger 3 [K]
T_{Sin}= Temperature of water at the inlet of solar heat exchanger [K]
T_{Sout}= Temperature of water at the outlet of solar heat exchanger [K]
W_{E1}= Work done by expander 1 [kW]
W_{E2}= Work done by expander 2 [kW]
$W_{expander}$= Total combined work done by the expanders [kW]

REFERENCES

[1] http://www.eia.gov/consumption/mahufacturing/index.cfm.(Cited by March 2016)
[2] Vélez, F., Segovia, J., Martín, M.C., Antolín, G., Chejne, F. and Quijano, A., "A technical, economical and market review of organic Rankine cycles for the conversion of low-grade heat for power generation". Renewable & Sustainable Energy Reviews, vol. 16, pp. 4175–4189, 2012.
[3] Vélez, F., Chejne, F., Antolín, G. and Quijano, "A., Theoretical analysis of a transcritical power cycle for power generation from waste energy at low temperature heat source".Energy Conversion and Management, vol. 60, pp. 188–195, 2012.
[4] Saleh, B., Koglbauer, G., Wendland, M. and Fischer, J., "Working fluids for low temperature organic Rankine cycles".Energy, Vol. 32, pp. 1210– 1221, 2007.
[5] Quolin, S., Declaye, S., Tchange, B.F. and Lemort, V., "Thermo-Economic optimization of waste heat recovery organic Rankine cycles". Applied Thermal Engineering, vol. 31, pp. 2885-2893, 2011.
[6] Tchange, B.F., Papadakis, G., Lambrinos, G. and Frangoudakis, A., "Fluid selection for a low-temperature solar organic Rankine cycle". Applied Thermal Engineering, vol. 29, pp. 2468–2476, 2009.
[7] Quolin, S., Aumann, R., Grill, A., Schuster, A., Lemort, V. and Spliethoff, H., "Dynamic modeling and optimal control strategy of waste heat recovery organic Rankine cycles". Applied Energy, vol. 88, pp. 2183– 2190, 2011.
[8] Hung, T.C., Shal, T.Y. and Wang, S.K.,"A review of organic Rankine cycles (ORC`s) for the recovery of low-grade waste heat". Energy, vol. 22 (7), pp. 661-667, 1997.
[9] Chen, H., Goswami, Y. and Stefanakos, E., "A review of thermodynamic cycles and working fluids for the conversion of low-grade heat". Renewable & Sustainable Energy Reviews, vol. 14, pp.3059–3067, 2010.
[10] Vélez, F., Segovia, J., Martín, M.C., Antolín, G., Chejne, F. and Quijano, A., "Comparative study of working fluids for a Rankine cycle operating at low temperature". Fuel Processing Technology, vol. 103, pp.71–77, 2012.
[11] U.S. Environmental Protection Agency. Class I Ozone Depleting Substances. [Online]. [date of reference March 11th of 2013] Available at: www.epa.gov/ozone/science/ods/classone.html.
[12] Roy, J.P., Mishra, M.K. and Misra, A., "Parametric optimization and performance analysis of a waste heat recovery system using organic Rankine cycle".Energy, vol. 35, pp. 5049-5062, 2010.
[13] Roy, J.P., Mishra, M.K. and Misra, A., "Parametric optimization and performance analysis of a regenerative organic Rankine cycle using low-grade waste heat for power generation". International Journal of Green Energy, vol. 8 (2), pp. 173–196, 2011.
[14] Manolakos, D., Papadakis, G., Mohamed, E., Kyritsis, S. and Bouzianas, K., "Design of an autonomous low-temperature solar Rankine cycle system for reverse osmosis desalination". Desalination, vol. 183, pp. 73–80, 2005.
[15] Manolakos, D., Papadakis, G., Kyritsis, S. and Bouzianas, K., "Experimental evaluation of an autonomous low-temperature solar Rankine cycle system for reverse osmosis desalination". Desalination, vol. 203, pp. 366–374, 2007.
[16] Manolakos, D., Kosmadakis, G., Kyritsis, S. and Papadakis, G., "On site experimental evaluation of a low-temperature solar organic Rankine cycle system for RO desalination". Solar Energy, vol. 83, pp. 646–656, 2009.
[17] Manolakos, D., Kosmadakis, G., Kyritsis, S. and Papadakis, G., "Identification of behaviour and evaluation of performance of small scale, low-temperature organic Rankine cycle system coupled with a RO desalination unit". Energy, vol. 34, pp. 767–774, 2009.
[18] Bruno, J.C., López-Villada, J., Letelier, E., Romera, S. and Coronas, A.,"Modelling and optimisation of solar organic Rankine cycle engines for reverse osmosis desalination". Applied Thermal Engineering, vol. 28, pp. 2212–2226, 2008.
[19] Delgado-Torres, A. and García-Rodríguez, L.,"Analysis and optimization of the low-temperature solar organic Rankine cycle (ORC)".Energy Conversion and Management, vol. 51, pp. 2846–2856, 2010.
[20] Thanche, B.F., Lambrinos, G., Frangoudakis, A. and Papadakis, G.,"Exergy analysis of micro-organic Rankine power cycles for a small scale solar driven reverse osmosis desalination system". Applied Energy, vol. 87, pp. 1295–1306, 2010.
[21] Schuster, A. and Karl, J.," Simulation of an innovative stand-alone solar desalination system using an organic Rankine cycle". Int. J. of Thermodynamics, vol. 10(4), pp. 155-163, 2007.
[22] Karellas, S., Terzis, K. and Manolakos, D., "Investigation of an autonomous hybrid solar thermal ORC-PV RO desalination system.

The Chalki island case." Renewable Energy, vol. 36, pp. 583-590, 2011.
[23] Kosmadakis, G., Manolakos, D. and Papakakis, G., "Parametric theoretical study of a two-stage solar organic Rankine cycle for RO desalination". Renewable Energy, 35, pp. 989–996, 2010.
[24] Franco, A. and Villani, M., "Optimal design of binary cycle power plants for water-dominated, medium-temperature geothermal fields". Geothermics, vol. 38, pp. 379–391, 2009.
[25] Aneke, M., Agnew, B. and Underwood, C., "Performance analysis of the Chena binary geothermal power plant".Applied Thermal Engineering, vol. 31, pp. 1825-1832, 2011.
[26] Astolfi, M., Xodo, L., Romano, M. and Macchi, E., "Technical and economical analysis of a solar–geothermal hybrid plant based on an organic Rankine cycle."Geothermics, vol. 40, pp. 58–68, 2011.
[27] Vaja, I. and Gambarotta, A., "Internal combustion engine (ICE) bottoming with organic Rankine cycles (ORCs)". Energy, vol. 34, pp. 767–774, 2009.
[28] Schuster, A., Karellas, S., Kakaras, E. and Spliethoff, H., "Energetic and economic investigation of organic Rankine cycle applications". Applied Thermal Engineering, vol. 29, pp. 1809–1817, 2009.
[29] Chen, H., Goswami, Y., Rahman, M. and Stefanakos, E., "A supercritical Rankine cycle using zeotropic mixture working fluids for the conversion of low-grade heat into power". Energy, vol. 36, pp. 549-555, 2011.
[30] Mikielewicz, D. and Mikielewicz, J., "A thermodynamic criterion for selection of working fluid for subcritical and supercritical domestic micro CHP".Applied Thermal Engineering, vol. 30, pp. 2357-2362, 2010.
[31] Lakew, A. and Bolland, O., "Working fluids for low temperature heat source". Applied Thermal Engineering, vol. 30, pp. 1262–1268, 2010.
[32] Lemmon, E.W., Huber, M.L. and Mclinden, M.O., Reference fluid thermodynamic and transport properties (REFPROP). NIST Standard Reference Database, 23, Version 8.0; 2007.
[33] FredyVélez, Farid Chejne & Ana Quijano "Thermodynamic analysis of R134a in an Organic Rankine Cycle for power generation from low temperature sources" Dyna, ISSN: 0012-7353,vol. 81, núm. 185, junio, 2014, pp. 153-159, Universidad Nacional de Colombia Medellín, Colombia.



Kaushalendra K. Dubey is associated with is the Dept of Mechanical Engineering,ShardaUniversity,Greater Noida,India as assistant professor. He is a research fellow of Delhi Technological University,Delhi,India. Author qualification is Master of technology in Thermal Engineering with honor's from Jamia Millia Islamia, Delhi, India. Currentl, He has faculty research fellowship from Indian Institute of Technology,Delhi in 2016 in Energy and Heat Recovery Research. Author has published more than 25 research papers in repute journals and conferences and active volunteer for work on energy research and technology through industrial and academic project.



Prof. (Dr.) R.S. Mishra Ph.D-SOLAR ENERGY IIT DELHI-1986., HOD Mechanical Engg in Delhi Technological University,Delhi,India . He has published more than 250 research papers in repute journals & conferences and also written 55 books on ASTROLOGY, VADIC Sciences, FENG SHUAI. Professor Mishra is dynamic personality of area of solar thermal system design-development and advanced refrigeration techniques.

Original Research Article **Microbiology**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

DEGRADATION OF PHORBOL ESTERS IN *JATROPHA* SEED CAKE BY *PSEUDOMONAS AERUGINOSA* DS1

SHILPI AHLUWALIA[1], JAI GOPAL SHARMA[1*] AND PUSHPENDRA SINGH[2]

[1]Department of Biotechnology, Delhi Technological University, Delhi- 110042, India.
[2]Department of Mechanical Engineering, Delhi Technological University, Delhi- 110042, India.

ABSTRACT

Jatropha curcas is a perennial plant which is widely used as biofuel and is recognized for its medicinal properties. Different parts of *Jatropha* plant are used for the treatment of human and veterinary diseases. Biodiesel production from *Jatropha* seed oil results in large amount of residual seed cake which is rich in lignocellulosic compounds, minerals and proteins but presence of certain toxic compounds limits the use of this nutrient rich cake. Phorbol esters are the major toxic compounds found in *Jatropha curcas* seed cake that make it unsuitable for food and feed applications. Effective utilization of this seed cake necessitates the need of complete removal of its phorbol esters. In this study, we investigated the application of submerged fermentation (SMF) for the detoxification of phorbol esters. SMF was carried out with *Jatropha* seed cake at 37 ˚C, pH 7, 100 rpm for 9, 12 and 15 h using *Pseudomonas aeruginosa* DS1. Complete degradation of phorbol esters resulted when the seed cake was fermented for 15 h at 37 ˚C, pH 7, 100 rpm indicating the potential of this process to completely detoxify the *Jatropha* seed cake in a very short time.

KEYWORDS: Detoxification, *Jatropha curcas*, Phorbol Esters, Submerged Fermentation

JAI GOPAL SHARMA*

Department of Biotechnology, Delhi Technological University, Delhi- 110042, India.

Received on: 18.01.2017
Revised and Accepted on: 24-03-2017
DOI: http://dx.doi.org/10.22376/ijpbs.2017.8.2.b542-546

INTRODUCTION

Jatropha curcas the wonder plant is drought resistant which has the potential to grow in marginal/poor soils. It has achieved tremendous interest because of its high oil content[1] which can be used for the production of biodiesel. Biodiesel can be considered as a substitute for diesel fuel as it gives a similar performance without modifications in the engine. At the same time, biodiesel production from *Jatropha* seeds leads to the production of large amounts of residual seed cake whose disposal or utilization should be guarded[2] which may be helpful in reducing the cost of *Jatropha* biodiesel industry. The *Jatropha* deoiled seed cake is rich in nutrients, but the presence of some anti-nutrients/toxins restricts the use of seed cake as animal feed. Phorbol esters have been identified as the major toxicants in seed cake as they have carcinogenic and mutagenic properties.[3] Many cases of *Jatropha* poisoning have been reported[4-6], but most of them go undocumented as symptoms are not specific and more of like bacterial or protozoal gastroenteritis which includes mainly abdominal pain, vomiting and diarrhoea. Various methods have been reported in the literature[7] to detoxify *Jatropha* seed cake which includes physical, chemical and radiation treatments but these methods require a consortium of treatments which makes the process expensive and time-consuming. Biological detoxification despite of the fact an environmentally and inexpensive process[8] is still at developing stage, although few researchers have reported the use of fungi, and bacteria for the detoxification of *Jatropha* seed cake by solid state fermentation process. The use of fungi for detoxification presented inadequate results as the metabolites produced by fungus are not effective in detoxifying complex phorbol esters.[9] Anjali and Haresh[10] succeeded in total removal of phorbol esters by solid state fermentation (SSF) of *Jatropha* seed cake with white rot fungi, but it took a long time of 20 days. Bacterial cultures may reduce the detoxification time which can make the process rapid and economical. 97% reduction in phorbol ester content was observed in 9 days when SSF of seed cake was done with *Streptomyces fimicarius.*[11] In another study, phorbol- ester content was reduced to an undetectable level when the seed cake was fermented (SSF) with *Pseudomonas aeruginosa* PseA strain within nine days at 30 ˚C, pH 7 and 65% relative humidity.[12] There was 93% reduction in the phorbol ester content when SSF was carried out by *Bacillus smithii* in 3 days.[13] Detoxification of phorbol esters by submerged fermentation process could be a good option than SSF as it requires less time with the process well suited for organisms requiring higher moisture levels for growth. This work aimed at detoxification of *Jatropha* seed cake by SMF using *Pseudomonas* sp. *Pseudomonas* has been well investigated as a lipase and protease producer[14] which may help in the degradation of toxic phorbol esters. SMF using *Pseudomonas* has never been reported in the literature for the detoxification of *Jatropha* seed cake.

MATERIALS AND METHODS

Materials and culture

Deoiled *Jatropha curcas* seed cake was obtained from Biodiesel Plant, DTU Delhi, India and it was ground and stored in ziplock pouches at 4˚C. The culture media were procured from HiMedia Laboratories (Mumbai, India). Standard, i.e., phorbol-12-myristate-13-acetate (Sigma-Aldrich, USA), HPLC grade acetonitrile (Fischer Scientific), HPLC grade methanol (SD fine, India) were used as received from the manufacturer. Other chemicals/solvents were of analytical grade and procured from standard companies. Strain *Pseudomonas aeruginosa* DS1 isolated from soil was used for degradation of phorbol esters.

Preparation of inoculums

Culture was grown overnight in nutrient broth at 37 ˚C and 100 rpm, and 1% of this culture was used to inoculate the media containing (%) starch 2; KH_2PO_4 0.5, KNO_3 1.01, NH_4Cl 0.535, $MgSO_4.7H_2O$ 0.001, $CaCl_2.2H_2O$ 0.01 and $Na_2HPO.12H_2O$ 0.8. The incubation was done in a rotary shaker at 37 ˚C for 24h at 100rpm and was centrifuged at 10000rpm for 20 min at 4˚C for obtaining a cell-free supernatant.

Submerged fermentation of deoiled Jatropha seed cake

Seed cake was autoclaved for 1 h and 5 g of it was added to the cell-free supernatant in different flasks. The flasks were incubated in a rotary shaker at 37 ˚C, pH 7, 100 rpm for 9, 12 and 15 h and phorbol esters were extracted. Control flasks containing 5 g seed cake in minimal medium (KH_2PO_4 0.5%, KNO_3 1.01%, NH_4Cl 0.535%, $MgSO_4.7H_2O$ 0.001%, $CaCl_2.2H_2O$ 0.01% and $Na_2HPO.12H_2O$ 0.8%) were also run alongside under same conditions. All the experiments were done in triplicates.

Extraction of phorbol esters from seed cake

Phorbol esters were extracted with absolute methanol by the method prescribed by Xing *et al.*[11] Briefly, methanol was added to the seed cake in 7:1 ratio and sonicated for 15 min and filtrate was collected. The extraction process was carried out thrice, and the extracted fractions were pooled together. Methanol was removed by rotary evaporator under vacuum at 60˚C; then the extract was fully dried in hot air oven at 40˚C.

Quantification of phorbol esters

Preliminary quantification of phorbol esters was done using thin layer chromatography. The untreated and treated seed cake sample extracts were column chromatographed over silica gel (60-120 mesh) and fractionated by eluting with 5% dichloromethane in acetonitrile. These fractions were analyzed for phorbol esters by TLC before confirmation analysis by HPLC (Agilent, 1260 series, USA) with UV detector (Agilent G1315D MWD). A reverse phase C18 column [5 µm, 4.6 X 250 mm i.d., Thermo Electron Corporation Massachusetts (MA) USA] was used in the method adopted for separation in accordance with Wink *et al.*[15] The solvent system used was water and acetonitrile [Starting with 60% water (A) and 40% acetonitrile (B) for 15 min, decreasing A to 25% and increasing B to 75% for next 20 min, and increasing B to 100% for next 10 min]. Separation was effected at room temperature

(25˚C) with a flow rate of 1.3 ml/min. PMA was used as a standard.

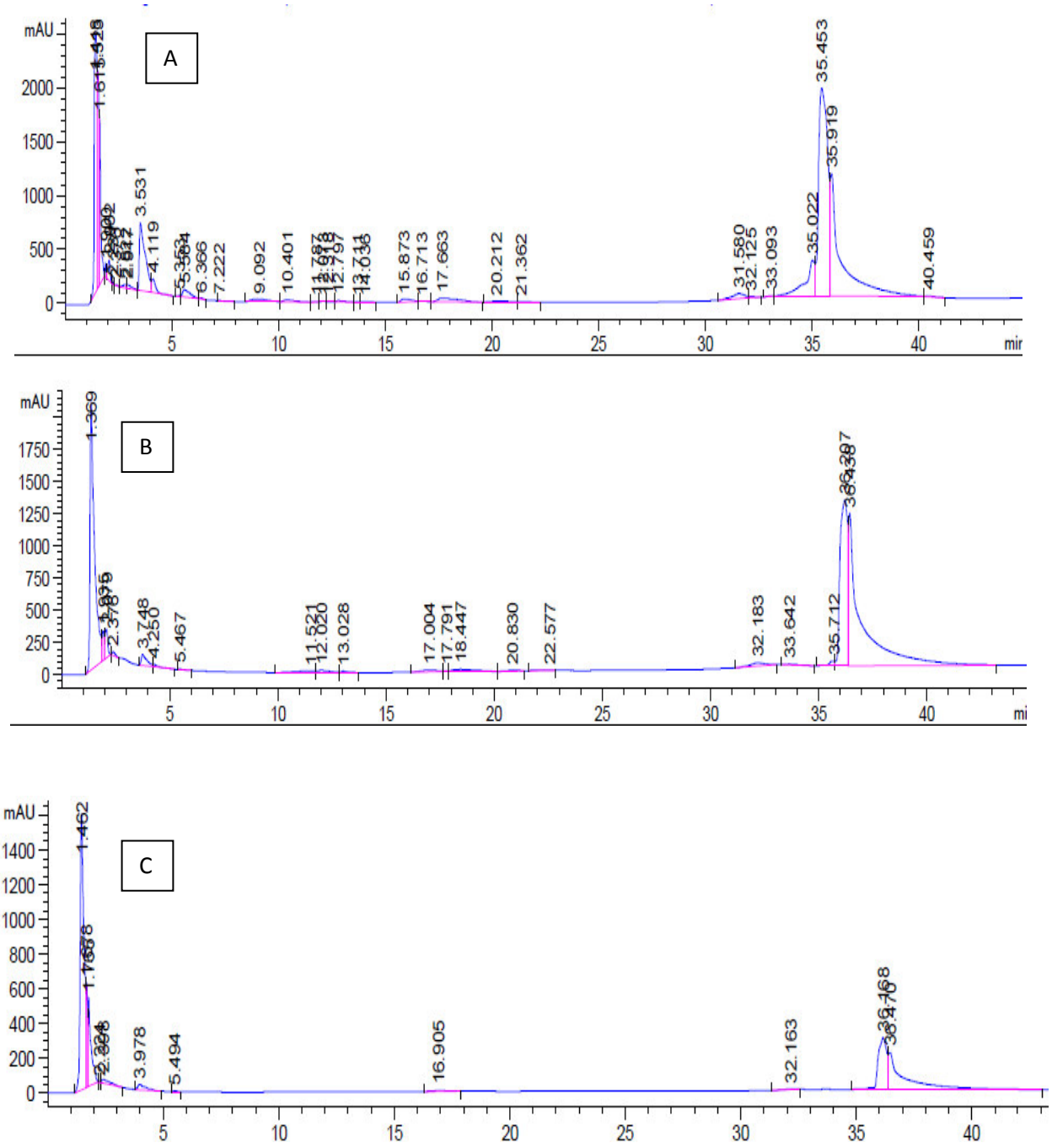
RESULTS

Phorbol esters in seed cake

Phorbol-12-myristate-13-acetate (PMA) is the only standard commercially available for investigating the presence of phorbol esters by HPLC. Methanol extracted phorbol esters from the deoiled seed cake were analyzed by HPLC. The quantification was done by taking PMA as the standard whose peak appeared at 37 min. HPLC chromatograms of *Jatropha* seed cake extract showed the presence of a major phorbol esters peak which consisted of two minor peaks at retention times of 35.4 and 35.9 min (Fig. 1A). We could identify two peaks for phorbol esters in the seed cake extracts. Total phorbol esters were calculated by summing up the area of phorbol ester peaks. The degradation was determined by comparing the area of the peaks for the samples with those of untreated seed cake.

Degradation of phorbol esters during submerged fermentation

Submerged fermentation was carried out using *Pseudomonas aeruginosa* DS1. The flasks containing the medium, seed cake and the inoculum were incubated in a rotary shaker at 37 ˚C, pH 7,100 rpm for 0, 9, 12 and 15 h and phorbol esters were extracted. Figure 1(A-D) shows the chromatograms of untreated and treated *Jatropha* seed cake samples at different time intervals. Phorbol ester concentration decreased gradually with increase in fermentation time. 31.57% reduction in phorbol ester peaks was observed after 9 h of fermentation period, which increased to 55.26% after 12 h. No phorbol ester peak was detected after 15 h of fermentation, and the phorbol ester content in the reference sample (uninoculated) remained same as that of 0 h.



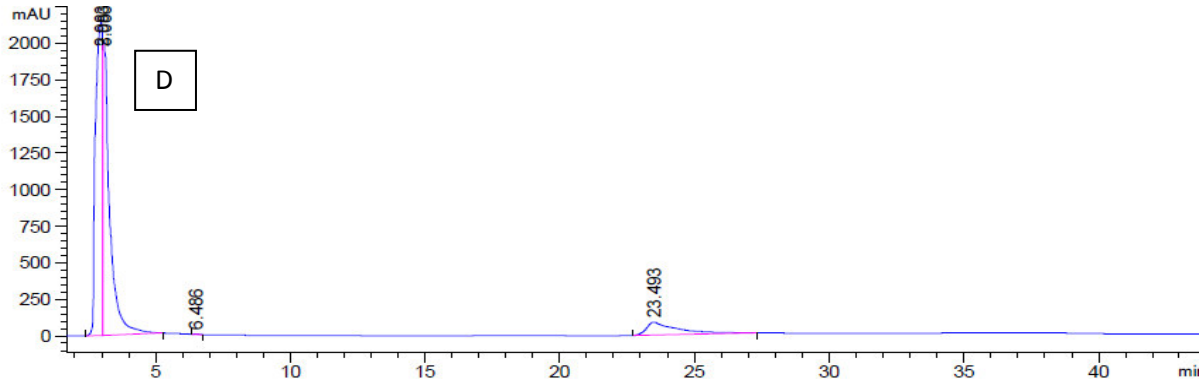


Figure 1

Chromatograms of untreated (A-for all sampling hours) and treated (B-9h; C-12h; D-15h) *Jatropha* seed cake samples.

DISCUSSION

Jatropha curcas contains six phorbol esters with diterpene moiety i.e. 12-deoxy-16-hydroxyphorbol in *Jatropha curcas* seed oil.[16] The quantity and the type of phorbol esters in *Jatropha* plant vary from country to country and is influenced by the species, soil fertility and the climatic conditions of the area where they grow.[16-18] In work done by Ahmed and Salimon [19], three different varieties of tropical *Jatropha curcas* were identified from Malaysia, Indonesia and India with two, five and four phorbol ester peaks respectively. We observed two phorbol esters peaks in our *Jatropha* seed cake sample which is similar to the findings of Chetna *et al.* [12] in which two peaks of phorbol esters were identified in the *Jatropha* seed cake, out of which the first peak with a retention time of 23.1 min was collected and quantified. Submerged fermentation technique has been the method of choice for bacterial enzyme production because of their higher moisture requirement. [20] In work done by Phengnuam and Suntornsuk[21] submerged fermentation resulted in more phorbol ester degradation in *Jatropha* seed cake as compared to solid state fermentation. While most of the research has been done with the biomass for detoxification, our work involved the use of enzymes in the broth after centrifugation. Degradation of phorbol esters was more when the inoculated seed cake was incubated for 15 h. Decrease in phorbol ester concentration with incubation time is attributed to certain enzymes like lipases and esterases released by the strain that hydrolyzes these diterpene esters.[22,23] In work done by Chetna *et al.*[12] lipase production by *Pseudomonas* was responsible for the possible degradation of phorbol ester during solid state fermentation of the *Jatropha* seed cake. So it can be inferred that the incubation time has a significant effect on the levels of phorbol esters of seed cake in the presence of biomass or free enzymes and our results clearly demonstrate that *Pseudomonas aeruginosa* DS1 could degrade phorbol esters in *Jatropha* seed cake to an undetectable level.

CONCLUSION

Submerged fermentation (SMF) was carried out to detoxify *Jatropha* seed cake by using *Pseudomonas aeruginosa* DS1. Complete degradation of phorbol esters was achieved when 5 g seed cake was inoculated at 37 ˚C, pH 7,100 rpm for 15 h which is very short treatment time compared to other methods of detoxification.The detoxification conditions may further be optimized to make the submerged fermentation process more suitable for commercial applications to produce animal feed supplements from *Jatropha curcas* seed cakes.

ACKNOWLEDGEMENT

The authors would like to thank Department of Chemistry, DTU for HPLC analysis. The first author acknowledge DTU for providing research fellowship.

CONFLICT OF INTEREST

Conflict of Interest declared none.

REFERENCES

1. Rathnakar Reddi KVN, Krishna Satya A, Ramesh P, Lekkala Sivaram P, Narendra K, Chandra ORP, Reddy CVCM, Chandra SA. Molecular genetic assessment of *Jatropha curcas* germplasm of diverse origin along with its wild relatives for various early growth and establishment related traits. *Int J Pharma Bio Sci*. 2016 ;7(4): 261 – 71.
2. Liang Y, Siddaramu T, Yesuf J, Sarkany N. Fermentable sugar release from *Jatropha* seed cakes following lime pretreatment and enzymatic hydrolysis. *Bioresour Technol.* 2010; 101: 6417–24.
3. Oskoueian E, Abdullah N, Ahmad S. Phorbol esters isolated from *Jatropha* meal induced apoptosis-mediated inhibition in proliferation of Chang and Vero cell lines. *Int J Mol Sci.*2012; 13: 13816–29.
4. Viral S, Jayesh S. Five cases of *Jatropha curcas* poisoning. *JAssoc Physic India.* 2010; 58: 243-44.
5. Singh RK, Singh D, Mahendrakar AG. *Jatropha* poisoning in children. *Med Jl Armed Forces India.* 2010; 66: 80-81.

6. Rai D, Lakhanpal P. *Jatropha curcus* poisoning in pediatric patients, Mauritius. *Internet J Pediatr Neonato.* 2007; 8(2): 946-50.
7. Ahluwalia S, Bidlan R, Sharma JG, Singh P. Review on phorbol ester degradation of *Jatropha* seed cake for its use as animal feed. *Intl JPharm Pharmal Sci.* 2017; 9(1): 7-13.
8. Rachna B, Renu D. Biological strategies for detoxification of hexavalent chromium. *Int J PharmaBio Sci.* 2017; 8(1): 35 – 48.
9. Belewu MA, Belewu KY, Ogunsola FO. Nutritive value of dietary fungi treated *Jatropha curcas* kernel cake: Voluntary intake, growth and digestibility coefficient of goat. *Agricult Biol JNorth Am.* 2010; 1(2): 135-38.
10. Anjali B, Haresh K. Phorbol ester degradation in *Jatropha* seed cake using white rot fungi. *Biotechnol.* 2014; 4: 447–50.
11. Xing HW, Lingcheng O, Liang-Liang F, Shui Z, Ji-Dong LGL, Jiao L, Changhe Z. Detoxification of *Jatropha curcas* kernel cake by a novel *Streptomyces fimicarius* strain. *JHazard Mater*. 2013; 260: 238– 46.
12. Chetna J, Priyanka M, Khare SK. Degradation of phorbol esters by *Pseudomonas aeruginosa* PseA during solid-state fermentation of deoiled *Jatropha curcas* seed cake. *Bioresour Technol.* 2011; 102(7): 4815-19.
13. Chang CF, Weng JH, Lin KY, Liu LY, Yang SS. Phorbol esters degradation and enzyme production by *Bacillus* using *Jatropha* Seed Cake as Substrate. *Int JEnviron Pollut Rem.* 2014; 2(1): 32-6.
14. Mahanta N, Gupta A, Khare SK. Production of protease and lipase by solvent tolerant *Pseudomonas aeruginosa* PseA in solid state fermentation using *Jatropha curcas* seed cake as substrate. *Bioresour Technol.* 2008; 99: 1729-35.
15. Wink M, Grimm C, Koschmieder C, Sporer F, Bergeot O. Sequestration of phorbol esters by the aposematically coloured bug Pachycoris klugii (Heteroptera: Scutelleridae) feeding on *Jatropha curcas* (Euphorbiaceae). *Chemoecology.* 2000; 10: 179–84.
16. Haas W, Sterk H, Mittelbach M. Novel 12-Deoxy-16-hydroxyphorbol diesters isolated from the seed oil of *Jatropha curcas*. *J Nat Prod.* 2002; 65: 1334-1440.
17. Martınez Herrera J, Siddhuraju P, Francis G, Davila-Ortiz G, Becker K. Chemical composition, toxic/antimetabolic constituents, and effects of different treatments on their levels, in four provenances of *Jatropha curcas* L. from Mexico. *Food Chem.* 2006; 96: 80–9.
18. Makkar HPS, Becker K, Sporer F, Wink M. Studies on nutritive potential and toxic constituents of different provenances of *Jatropha curcas*. *J AgricultFood Chem*. 1997; 45: 3152-7.
19. Ahmed WA, Salimon J. Phorbol esters as constituents of tropical *Jatropha curcas* seed oil. *Eur J Sci Res.* 2009; 31: 429–36.
20. Subramaniyam R, Vimala R. Solid state and submerged fermentation for the production of bioactive substances: A comparative study. *Int JSciNat.* 2012; 3(3): 480-6.
21. Phengnuam T, Suntornsuk W. Detoxification and anti-nutrients reduction of *Jatropha curcas* seed cake by *Bacillus* fermentation.*JBiosciBioeng.* 2013; 115(2): 168-72.
22. De Barros CRM, Ferreira LMM, Nunes FM, Bezerra RMF, Dias AA, Guedes CV, Cone JW, Marques GSM, Rodrigues MAM. The potential of white-rot fungi to degrade phorbol esters of *Jatropha curcas* L. seed cake.*Eng Life Sci.* 2011; 11: 107–10.
23. Azhar N, Norhani A, Wan Z, Syahida A, Ehsan O, Faridah A, Youssuf G. Detoxification of Toxic Phorbol Esters from Malaysian *Jatropha curcas* Linn. kernel by *Trichoderma* spp. and Endophytic Fungi. *Int J Mol Sci.* 2014; 15: 2274-88.

Reviewers of this article

Dr. Gaganpreet Kaur Saini

Doctorate in Biotechnology, FSQA Coordinator,Kiskadee Ventures Limited,130 Dynamic Drive,Scarborough Canada,M1V 5C9



Prof. Y. Prapurna Chandra Rao

Assistant Proffessor, KLE University, Belgaum, Karnataka



Prof.Dr.K.Suriaprabha

Asst. Editor , International Journal of Pharma and Bio sciences.



Prof.P.Muthuprasanna

Managing Editor , International Journal of Pharma and Bio sciences.

We sincerely thank the above reviewers for peer reviewing the manuscript

Environ Dev Sustain
DOI 10.1007/s10668-017-9981-1

Determination of hierarchical relationships among sustainable development goals using interpretive structural modeling

Pravin Kumar[1] · Faisal Ahmed[2] · Rajesh Kumar Singh[3] · Prerna Sinha[4]

Received: 1 December 2016 / Accepted: 28 May 2017
© Springer Science+Business Media Dordrecht 2017

Abstract This paper analyses the complex interactions among the seventeen sustainable development goals (SDGs) and presents them in a hierarchical framework using interpretive structural modeling technique. The driving and dependence powers of each of these goals have been ascertained to help the developing and least developed countries improve their strategic orientation for a particular goal within a given time frame. This would also help them in prioritizing their resource allocations on a specific SDG by focusing on its hierarchical ranking on one hand and driving and dependence powers on the other.

Keywords SDG · MDG · Sustainable development · Developing countries · Interpretive structural modeling

✉ Pravin Kumar
pravin.dce@gmail.com

Faisal Ahmed
f_faisal09@yahoo.co.in

Rajesh Kumar Singh
rksdce@yahoo.com

Prerna Sinha
dr.prerna.sinha@gmail.com

[1] Department of Mechanical Engineering, Delhi Technological University, Bawana Road, Delhi 110042, India

[2] FORE School of Management, Delhi 110016, India

[3] Department of Operations Management, Management Development Institute, Gurgaon, Haryana, India

[4] Department of Geography, Tilkamanjhi Bhagalpur University, Bhagalpur, Bihar, India

Published online: 05 June 2017

1 Introduction

The idea of sustainable development is a not a new phenomenon, but has maintained its continuity and relevance in a globalizing world. In September 2015, the United Nations General Assembly adopted the post-2015 development agenda entitled "Transforming our world: the 2030 Agenda for Sustainable Development." The preamble of this document calls for strengthening universal peace and for the eradication of poverty to set forth the people and the planet on a path of sustainable development. It says "eradicating poverty in all its forms and dimensions, including extreme poverty, is the greatest global challenge and an indispensable requirement for sustainable development" (UNGA 2015).

The 2030 agenda includes 17 goals, which are referred to as sustainable development goals (SDGs) and a total of 169 targets are integrated with these goals. The call for an initiative for the development of SDGs was a key outcome of the Rio+20 Summit held in 2012. The goals and their targets have been prepared duly considering the country's level of development and their capacity constraints. It is, therefore, imperative that to achieve the best result by 2030, these goals and targets must be aligned with a country's national development plan. SDGs, in fact, are a successor of the millennium development goals (MDGs), adopted by world leaders in the year 2000. MDGs identified eight well-defined and time-bound goals to address rudimentary global concerns like eradication of extreme poverty and hunger, achieving universal primary education, reducing child mortality, among others. They were to be achieved by 2015 but many countries fell short of their targets, thus creating the need for a post-2015 developmental agenda in the form of SDGs. There are some academic studies on MDGs, e.g., Haines and Cassels (2004), Waage et al. (2010), Sachs (2012) and Esquivel and Sweetman (2016), which have been discussed later in this paper. However, the present paper focuses on developing a hierarchical model of SDGs through a modeling technique to enable its use in policy making.

Since the SDGs were implemented only recently in 2015, therefore there has been limited use of modeling technique in analyzing them. One study by Joshi et al. (2015) focused on governance issues for post-2015 SDGs and forecasted the future of domestic governance through the year 2060 for 183 countries. They use a long-term, dynamic, integrated global futures modeling system. However, this study does not take into account each of the SDGs, as they were finalized only in the later half of 2015. Thus, we primarily cite only one study, i.e., Costanza et al. (2016a), which uses a modeling approach to establish a relationship among the SDGs. They developed a sustainable well-being index (SWI) that is a function of net economic contribution, natural capital/ecosystem services contribution, and social capital/community contribution. Further, they used a hybrid approach that proposes to integrate SDGs and SWI into a nonlinear, system dynamics model capable of presenting various policy scenarios for the future. Therefore, there is a need to study the complex interconnections among the goals (Costanza et al. 2016a; Costanza et al. 2014). All these 17 goals cannot function in a mutually exclusive manner because each one may influence few others and in turn be influenced by them as well. This implies that there is a marked cause and effect relationship among the goals which cannot be undermined. In this paper, we assess the interactive relationships among them using interpretive structural modeling (ISM), a widely used mathematical modeling approach. This would be helpful in understanding their mutual relationship and how they influence the fulfillment of SDGs. Griggs et al. (2014) also pointed that in terms of targets, arguing that some targets should deliberately address interactions, thus creating avenues for synergies and trade-offs.

We, thus, need to ascertain what could be the driving power of each goal, so that the countries can prioritize their developmental goals and enhance their preparedness in terms of resource allocations for different sectors. For prioritizing the SDGs, a systematic method could be to present it in a hierarchical form so that the countries can identify the challenges and deterrents as identified by Waage et al. (2010) in the case of the MDGs. Though most of these deterrents seem to have been considered while framing SDGs, yet putting these goals into a hierarchical framework will bring ease to the countries in prioritizing the resource allocations for the goals. It would help countries align the SDGs with the national development plans or design their future development plans in line with their commitments to SDGs. Therefore, we have used interpretive structural modeling (ISM) to find the driving power and dependence of the 17 SDGs, and place them in a hierarchical framework.

Further, this paper is arranged into four sections. The first section examines the thematic discourse; the second section presents the methodology; the third section discusses the analysis and outcome of the research, and finally, the fourth section of the paper concludes the research work.

2 Literature review

The environmental concerns of economic growth and the need for sustainable development have garnered enough global attention in academic and policy discourses in the last few decades. In 1987, the Brundtland Commission set up by the United Nations presented its report entitled "Our Common Future" and explained what sustainable development implied for preserving the environment and natural resources. The report defined sustainable development as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (WCED 1987).

However, the issue of sustainable development got its stimulus from the United Nations Conference on Environment and Development popularly known as the Earth Summit or the Rio Summit of 1992. It factored-in a strong global realization for the cause and garnered enough international attention needed to bring the sustainability agenda at the core of policy discourse in the 1990s and thereafter. Studies have also focused on the prospects and implementation paradigms of one or more aspects of the Earth Summit outcomes (Cicin-Sain et al. 1995).

Later, in September 2000, the world leaders adopted the United Nations Millennium Declaration and identified eight goals, known as the MDGs. Each goal comprised of quantified targets and indicators thereof. The developed, developing, and the least developed countries (LDCs) have cooperated toward ensuring sustainable livelihoods for the present and the future generations. Sachs (2012) has rightly termed MDGs as a historic method of global mobilization aimed to achieve a set of important social priorities. In the foreword of the MDG Report 2015, the UN Secretary General Ban Ki-Moon writes that "The MDGs helped to lift more than one billion people out of extreme poverty, to make inroads against hunger, to enable more girls to attend school than ever before and to protect our planet. They generated new and innovative partnerships, galvanized public opinion, and showed the immense value of setting ambitious goals" (MDG Report 2015).

Haines and Cassels (2004) studied the progress of MDGs across different countries and regions. Though the countries strived hard to achieve their MDG targets by 2015, yet there have been some gaps which were well reflected in the MDG Report 2015, in terms of

attainment of the goals and their targets. Waage et al. (2010) made a cross-sectional analysis of the MDG and identified the challenges that were deterring its achievements. Some of the identified challenges include reducing gaps and fragmentation in the MDGs, improving targets and indicators, and ensuring equity and ownership. Later, in 2012, Rio+20 Summit, which took a stock of the situation after 20 years of the Earth Summit, also agreed that the aims of sustainable development goals (SDGs) should focus on alleviating poverty and creating a support system which is stable and resilient. SDGs thus are the outcomes of the urge for a more pragmatic development agenda post-2015. When the SDGs are compared to MDGs, definitely the new agenda and goals are well designed and relatively better (Esquivel and Sweetman 2016). SDGs are a successor of the development pursuits of the MDGs, but with potentially larger global aspirations and a relatively better national preparedness. With a sense of urgency, the new 17 goals are now forcing the international community and national governments alike to address their developmental needs in a way that also creates sufficient opportunities for the future generations to meet their own needs.

In 2012, the Rio+20 Summit called on the nations to develop a set of universal sustainable development goals (SDGs). These new goals were a preparation of the post-2015 agenda and would succeed the MDGs (United Nation 2012). The 2030 agenda includes 17 goals, which are referred to as sustainable development goals (SDGs) and a total of 169 targets and 300 indicators integrated with these goals (UNGA 2015). These goals are—Goal 1: No poverty; Goal 2: Zero hunger; Goal 3: Good health and well-being; Goal 4: Quality education; Goal 5: Gender equality; Goal 6: Clean water and sanitation; Goal 7: Affordable and clean energy; Goal 8: Decent work and economic growth; Goal 9: Industry, innovation and infrastructure; Goal 10: Reduced inequalities; Goal 11: Sustainable cities and communities; Goal 12: Responsible consumption and production; Goal 13: Climate action; Goal 14: Life below water; Goal: 15: Life on land; Goal 16: Peace, justice and strong institutions; and Goal 17: Partnerships for the goals. These goals and their targets, which amalgamate the developmental agenda with capacity constraints, clearly signify there is a need to reevaluate the relationship between people and planet for sustainability. The overarching aims of the SDGs, as agreed by nations at Rio+20, can be summarized as poverty eradication, sustainable lifestyles for all, and a stable resilient planetary life-support system. However, it is challenging to create and agree on SDGs that meet these overarching aims while resolving potential interactions among the sectoral goals.

SDGs offer a comprehensive dashboard of goals, targets, and indicators. SDGs address some of the systemic barriers to sustainable development and contain better coverage of, and balance between, the three dimensions of sustainable development viz. social, economic, and environmental. It is also said that "eradicating poverty in all its forms and dimensions, including extreme poverty, is the greatest global challenge and an indispensable requirement for sustainable development" (UNGA 2015). In the year 2000, the United Nations General Assembly (UNGA) adopted the Millennium Declaration to address the rudimentary challenges of extreme poverty and hunger, and child and maternal health, among others. It set out goals with time-bound targets to be attained through a commitment to global partnership. These goals are popularly known as the Millennium Development Goals (MDGs). It comprised of 8 goals, 21 targets, and 60 indicators to be achieved by the end of 2015. In fact, one of the main outcomes of Rio+20 was an international agreement to negotiate a new set of SDGs as a post-2015 development agenda.

The existing literature does not point out any possible balance between the various social, economic, and environmental objectives for sustainable development. Sachs (2012) clearly asserts that till now there is no consensus regarding the trade-offs and synergies

across the economic, environmental, and social objectives. The trade-offs between population, production, and consumption often lead to conflict between socioeconomic development and global sustainability (Griggs et al. 2014). Griggs et al. (2014) developed a framework of goals and their associated targets and argued that it was possible to develop integrated targets pertinent to various economic, social, and environmental concerns. Although they tried to establish the interactions between social and biophysical targets, yet the focus has not been on goals, as this study drew inferences from Griggs et al. (2013) who proposed six SDGs, discussed later in this paper.

Before the SDGs were to be adopted by the United Nations, Norström et al. (2014) advocated that the process of formulation of SDGs should be based on three considerations: (a) embracing an integrated social-ecological system perspective, (b) addressing trade-offs between the goals and the feasibility of achieving them considering biophysical, social, and political constraints, and (c) the exercise of goal-setting and its implementation should be considered the process of social change at individual as well as global levels. After the SDGs were proposed, very few studies have been done considering either few or all SDGs. Costanza et al. (2016a) clustered all the 17 SDGs into three elements of sustainable well-being namely (a) efficient allocation: building a living economy (SDGs 7, 8, 9, 11, 12); (b) fair distribution: protecting capabilities for flourishing (SDGs 1, 2, 3, 4, 5, 10, 16.17); and (c) sustainable scale: staying within planetary boundaries (SDGs 6, 13, 14, 15). They also proposed the concept of SWI as explained in the previous section of this paper.

The most recent studies published in 2016 point out to goal-specific orientations and thus contribute to the thematic discourse. For instance, Esquivel and Sweetman (2016) have focused on Goal 5, i.e., gender equality. They make a comparison between the strength of MDGs and SDGs to contribute to the cause of gender equality. They have argued that in the non-gender MDGs, the gender analysis was weak or nonexistent; but many gender issues have been well integrated even in the "non-gender" SDGs. A similar argument has been put forth by Fredman et al. (2016) who mention that gender equality is mainstreamed into various other SDGs and carries human rights-based orientation. They have appraised the inclusion of nine specific targets therein including those related to eliminating gender-based violence and access to reproductive health.

Also, considering another dimension viz. financial inclusion, Klapper et al. (2016) presented a case of the need for financial inclusion as a means to fulfill some of the SDGs. They argue that the access to financial services can help in ending poverty or providing quality education or the financial inclusion of farmers can lead to a higher degree of food security. However, they have not mentioned any relations between financial inclusion and other SDGs, arguing that they are not as well established.

In 2002, the World Summit on Sustainable Development in Johannesburg came up with five priority areas: water and sanitation, biodiversity and ecosystem management, energy, agricultural productivity and health (Annan 2002; Bouma 2014). A set of six SDGs were listed by augmenting MDGs with necessary conditions to ensure the stability of Earth's systems (Griggs et al. 2013). These are listed as follows:

Goal 1	Thriving lives and livelihoods: including poverty alleviation and improved well-being
Goal 2	Sustainable food security: including ending hunger
Goal 3	Sustainable water security: including universal access to clean water and sanitation
Goal 4	Universal clean energy: including a focus on climate change mitigation

Goal 5	Healthy and productive ecosystems: including focus on sustaining biodiversity
Goal 6	Governance for sustainable societies: including a focus on transform governance and institutions

Further, some of the studies include proposing an SDG for health based on mortality trends (Norheim et al. 2015); proposing an SDG indicator for global wastewater treatment (Malik et al. 2015); presenting a case for including mental health as a SDG (Thornicroft and Patel 2014); advocating the need for a mental health target and indicator (Votruba et al. 2014); understanding the transformational challenges in developed countries in context of universal SDGs (Osborn et al. 2015); assessing the impact of SDGs on the European sustainable development governance framework (Pisano et al. 2015); finding out how SDGs offers new opportunities for the Tropical Delta Regions (Szabo et al. 2015); and ascertaining the need for institutional restructuring while shifting from MDGs to SDGs (Kamepalli and Pattanayak 2015).

On the specific SDGs, Pablos-Mendez et al. (2016) highlighted the inclusion of universal health coverage (UHC) as a target in SDG 3, which is related to health. However, they have also argued that there is no standard prescription and a higher political commitment or societal value is needed to achieve it. They also call for domestic resource mobilization to enhance organized partnerships from public and private sectors in order to ensure a public health facility.

Osborn et al. (2015) examined how the SDGs pose a challenge to developed countries as well in the context of transforming their own societies as well as in contributing to the global efforts in developing countries. The developed countries have been assessed on an 8-point scale. Based on these scores, the priority rank of the SDGs has been changed to reflect a developed country context. The highest scoring goals appear at the top and depict the biggest transformational challenge. The goals are then ranked in a descending order, with goals with the lowest score appearing at the bottom. For instance, they observe that Target 1.2 of Goal 1, i.e., "No poverty," can refer to a particular kind of poverty still found in developed countries. This can include challenges pertaining to poverty among children, elderly minority, etc., and can even include homelessness or food poverty. Osborn and Ullah (2014) have also questioned the proposed SDGs as being too much focused on developing countries' role. They argue that the SDGs ignored the need for sustainability transformation in developed and middle-income economies, phenomena which would consequently affect the developing countries (Pisano et al. 2015).

According to Stokstad (2015), only 29% of the 169 targets are well-defined and scientifically rigorous. The others lack specific endpoints and time frames, and some cannot be measured accurately. For example, the target for water and sanitation is ambiguous and does not specify what percentage of a population should experience improvements by 2030. Such undefined or vague targets could deter the participation of donor agencies. They suggest to reduce the number of people suffering from water scarcity. A current target is to halve the number of sufferers; however, the access is not defined. The report says that only 29% of SDG's targets are well-defined, while 54% need work and 17% are weak or nonessential.

To guide this governance effort, five key principles have been outlined in the documents of the UN and the European Union (Berger 2009; Steurer 2009; Lange et al. 2013): horizontal policy integration, vertical policy integration, participation, reflexivity and the integration of long- and short-term time scales. Niestroy (2014) has compiled and conceptualized these principles to depict specific governance challenges for sustainable

development (Table 1). Joshi et al. (2015) also focused on the issues of security, capacity building and inclusion for governance as important governance issues for SDGs.

Some of the important previous studies done on SDGs are summarized in Table 2.

There are three elements that cover the usual three components of sustainable development viz. economy (implying the quality of life and well-being), society (implying equitable distribution), and environment (implying socio-symbiotic relationship) (Costanza et al. 2014). Esquivel (2016) offers a power analysis of SDGs process and outcomes from a feminist perspective. Many see, in the SDGs, several opportunities for progress on gender equality and women's rights, if not for transformation. Yet there are many reasons for skepticism, as the vision of the 2030 Agenda may not merely be attained through the use of rhetoric or highlighting funding provisions.

Therefore, to achieve the SDGs, policymakers, scientists, and practitioners ought to clarify how the goals and targets interconnect, including trade-offs and synergies, and develop three additional elements: (a) an aggregation of metrics of human and ecosystem well-being, (b) dynamic models of the integrated system of humans and the natural world, and (c) innovative ways to build broad public consensus on the future we want—the details of a world in which the SDGs have been implemented (Costanza et al. 2016b).

The obsessions with gross domestic product (GDP) growth rate have led to inequality and environmental damage in many countries. If one takes these elements into account, as the Genuine Progress Indicator (GPI) does, then arguably there has been no net progress globally for decades. Increasing income inequality, environmental damage, and other costs can cancel out positive gains from GDP growth (Costanza et al. 2016b).

3 Research methodology

This paper incorporates experts' opinions to decipher the relationships among the sustainable development goals and uses interpretive structural modeling (ISM) to develop the hierarchical relationship. An open discussion was done with a number of experts engaged

Table 1 Normative governance principles for sustainable development (Steurer 2009; Niestroy 2014)

Characteristics of Integration	Governance principle	What needs to be coordinated/ integrated?	Challenge for governance for SD
Policy sectors/ areas	Horizontal coordination/ integration	Multiple sectors: economic, social, and environmental policies	Policy coherence and systematic thinking
Policy scales/ spatial scales	Vertical coordination/ integration	Multiple levels: local, sub-national, national, super-national	Coordination. Links and cooperation of governments and administrations at all levels
Actors/ Stakeholders	Participation	Multiple actors: decision makers and stakeholders from politics, business, and civil society	Stakeholder participation in discussion and decision-making processes
Knowledge	Reflexivity	Knowledge and experience from various areas in society, learning the process	Continuous reflection and evaluation of decisions and policy instruments
Time	Intra- and intergenerational equity	Long- and short-term thinking	Long-term thinking in spite of short election cycles

Table 2 Summary of past works addressing the 17 goals of sustainable development

References	Goal 1: No poverty	Goal 2: Zero hunger	Goal 3: Good health and well-being	Goal 4: Quality education	Goal 5: Gender equality	Goal 6: Clean water and sanitation	Goal 7: Affordable and clean energy	Goal 8: Decent work and economic growth	Goal 9: Industry, innovation and infrastructure
Costanza et al. (2016a, b)	√	√	√	√	√	√	√	√	√
UNGA (2015)	√	√	√	√	√	√	√	√	√
Norheim et al. (2015)			√						
Szabo et al. (2015)									√
Esquivel and Sweetman (2016)					√				
Pablos-Mendez et al. (2016)			√						
Kamepalli and Pattanayak (2015)	√	√	√	√	√	√	√	√	√
Bouma (2014)									
Annan (2002)		√	√			√	√		
Pinter et al. (2013)	√	√	√	√		√	√	√	√
Akenji and Bengtsson (2014)		√					√		
Fredman et al. (2016)					√				
UNGA (2015)	√	√	√	√	√	√	√	√	√
Griggs et al. (2013, 2014)	√	√	√	√	√	√	√	√	√

Table 2 continued

References	Goal 10: Reduced inequalities	Goal 11: Sustainable cities and communities	Goal 12: Sustainable consumption and production	Goal 13: Climate action	Goal 14: Life below water	Goal 15: Life on land	Goal 16: Peace, justice and strong institutions	Goal 17: Partnership for the goals
Costanza et al. (2016a, b)	√	√	√	√	√	√	√	√
UNGA (2015)	√	√	√	√	√	√	√	√
Norheim et al. (2015)								
Szabo et al. (2015)				√	√	√		
Esquivel and Sweetman (2016)								
Pablos-Mendez et al. (2016)								
Kamepalli and Pattanayak (2015)	√	√	√	√	√	√	√	√
Bouma (2014)				√		√		
Annan (2002)				√	√	√		
Pinter et al. (2013)	√		√	√	√	√		√
Akenji and Bengtsson (2014)			√	√	√	√		
Fredman et al. (2016)								
UNGA (2015)	√	√	√	√	√	√	√	√
Griggs et al. (2013, 2014)	√	√	√	√	√	√	√	√

in research and consultancy work and also those who are supervising research in related domains. Broadly, we have divided the SDGs into three main domains—environmental sustainability goals, economic sustainability goals, and the social sustainability goals. The name and affiliation of these experts have been acknowledged in this paper. They were asked to give their views on linkages and relationships among various SGDs. Their opinions were noted, and most common opinion has been incorporated in the self-interaction matrix of ISM. The linkages among the SDGs are subjective, and different aspects of these relationships were examined in detail by these experts.

ISM is a technique which presents specific relationships among the variables and their hierarchical structure in a directed graph model. Primarily, it is a group learning process, but can be used individually as well. This model was first employed by Mandal and Deshmukh (1994) in their study on vendor selection, and since then has been used frequently in different contexts.

ISM technique includes the following steps:

Step 1 Identifying relevant enablers through group problem-solving technique

In this paper, all the 17 SDGs are used as enablers.

Step 2 Establishing mutual relationship among enablers

The relationships among the enablers may be established by individual decision or group decision-making process. In this paper, a number of experts are consulted and their opinions regarding the relationships are incorporated.

Step 3 Developing a structural self-interaction matrix (SSIM) of enablers showing a pair-wise relationship

In this paper, these relationships are shown in terms of V, A, X, and O as discussed in the next section.

Step 4 Preparing an initial reachability matrix from the SSIM, and based on transitivity, prepare the final reachability matrix

An underlying assumption herein is that if enabler A influences enabler B, and B influences enabler C, then A will necessarily influence C. Thus, the indirect relationships are observed using this step.

Step 5 Using iterations to rank the enablers at various levels

Table 6 is a summary table of the level partition which is carried out on iteration basis. In the first level, those goals are selected whose reachability set and the interaction set are exactly same, e.g., in the first iteration, we can see that Goal 17 has the same reachability and interaction set. Therefore, Goal 17 is kept on the first level. Therefore, the digit indicating 17 has been removed from all the rows. We repeat the iteration and find that in the second iteration, Goal 16 and Goal 3 have same reachability and interaction sets. Thus, both the Goals 16 and 3 are kept at the second level. Again the digit indicating 16 and 3 is removed from all the rows, and the process is repeated till all the goals are leveled.

Step 6 Prepare a digraph based on final reachability matrix;

After level partition and simplest digraph is prepared in the form of the hierarchy of relationships among the sustainable development goals.

Step 7 The digraph is then converted into an ISM-based model by replacing element nodes with the statements;
Step 8 Checking the model for conceptual inconsistency, and modifying, if need be

4 Analysis and findings

After finalizing the relationship among the various SDGs, the structural self-interaction matrix is developed and a hierarchy of the SDGs is prepared using all the steps of ISM in the following manner:

4.1 Structural self-interaction matrix

A contextual relationship of depicting whether one enabler "influences" the other is used. In our paper, to develop these contextual relationships among the 17 SDGs, expert opinions based on brainstorming has been used. In this study, the experts consulted are academics from economics, environmental science, and sociology disciplines. For expressing the relationship among different goals to achieve sustainable development, we have used four symbols to specify the direction of the relationship between any two SDGs, say i and j as shown in Table 3. These directions are stated as follows:

(a) V: goal i influence goal j;
(b) A: goal j influence goal i;
(c) X: goal i and j will influence each other; and
(d) O: goals i and j have no relationship.

5 Reachability matrix

The initial reachability matrix is prepared by converting SSIM through binary notations. This is done by substituting V, A, X, and O by 1 and 0 according to specified rules, mentioned as follows:

(a) If (i, j) entry in SSIM is V, the (i, j) entry in the reachability matrix is assigned 1, whereas the (j, i) entry is assigned 0;
(b) Similarly, if the (i, j) entry is A, it will be assigned 0 and 1, respectively;
(c) Similarly, if the (i, j) entry is X, it will be assigned as 1 in both cases.
(d) Similarly, if the (i, j) entry is O, it will be assigned as zero in both cases.

where "i" indicates the row and "j" indicates the column.

Table 4 shows the initial reachability matrix for the SDGs. After incorporating the transitivity as described earlier in Step 4 of ISM technique, the final reachability matrix is prepared as shown in Table 5. This matrix shows the driving power and dependence for each of the goals. The driving power of an SDG implies the total number of SDGs (including itself) which it may help to achieve. However, dependence power implies the total number of goals (including itself), which may help to achieve it. The driving and dependence power are used to classify the enablers, i.e., SDGs in this case, into the four groups viz. autonomous, dependent, linkage, and independent barriers in a specialized analysis called MICMAC (Impact Matrix Cross-Reference Multiplication Applied to a

Table 3 Structural self-interaction matrix (SSIM)

Sustainable development goals (SDGs)[a] $i \downarrow \rightarrow j$	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2
No poverty	V	O	O	O	O	O	O	O	O	A	O	O	O	O	V	V
Zero hunger	V	O	O	O	O	O	O	O	O	O	O	O	O	O	V	
Good health and well-being	V	O	O	O	O	O	O	O	O	O	O	O	O	O		
Quality education	O	O	O	O	V	O	O	V	O	V	O	O	V			
Gender equality	O	O	O	O	O	O	O	O	O	V	O	O				
Clean water and sanitation	O	O	V	V	A	O	V	O	O	O	A					
Affordable and clean energy	O	O	V	V	A	V	V	O	V	V						
Decent work and economic growth	O	V	O	O	O	A	A	A	A							
Industry, innovation, and infrastructure	O	O	O	O	O	O	V	O								
Reduced inequalities	O	O	O	O	O	O	O									
Sustainable cities and communities	O	O	A	A	O	A										
Responsible consumption and production	O	O	A	A	O											
Climate action	O	O	V	V												
Life below water	O	O	X													
Life on land	O	O														
Peace, justice and strong institutions	V															
Partnerships for the goals																

[a] i indicates the row and j indicates the column

Classification). The MICMAC analysis is a structural analysis which helps to graphically depict the driving and dependence power of enablers.

6 Level partitions

Based on the final reachability matrix, the reachability and antecedent sets for each of the goals are found. The reachability set comprises of the goal itself and other goals which it may help achieve, whereas the antecedent set consists of the goal itself and the other goals, which may help to achieve it. Then, the intersection of these sets is derived for all the goals. Level partition is done manually on iteration basis as earlier explained in Step 6 of the ISM technique.

These identified levels help in building the directed graph and the final hierarchical model. Table 6 shows that the SDG no. 17, i.e., "Partnerships for the Goals," is placed at level I. Thus, it would be positioned at the top of the ISM hierarchy. This iteration is repeated to find out the level of each goal as shown in Table 6. Based on the identification of level for each goal, the final hierarchical model of ISM is prepared. Figure 1 shows levels of all the goals in a digraph.

It is found that the provision of Quality Education (Goal 4) is the main driver to achieve the other SDGs. Quality education directly influences the awareness of the people about

Table 4 Initial reachability matrix

SDGs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	1	1
6	1	1	1	0	0	1	0	1	0	0	1	1	0	1	1	1	1
7	1	1	1	0	0	1	1	1	1	0	1	1	0	1	1	1	1
8	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	1	1
9	1	1	1	0	0	0	0	1	1	0	1	0	0	0	0	1	1
10	1	1	1	0	0	0	0	1	0	1	0	0	0	0	0	1	1
11	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0	1	1
12	1	1	1	0	0	0	0	1	0	0	1	1	0	0	0	1	1
13	1	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1
14	1	1	1	0	0	0	0	1	0	0	1	1	0	1	1	1	1
15	1	1	1	0	0	0	0	1	0	0	1	1	0	1	1	1	1
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table 5 Final reachability matrix after incorporating transitivity

SDGs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Driving power
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	**4**
2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	**3**
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	**2**
4	***1***	***1***	***1***	1	1	***1***	***1***	1	***1***	1	***1***	***1***	1	***1***	***1***	***1***	***1***	**17**
5	***1***	***1***	***1***	0	1	0	0	1	0	0	0	0	0	0	0	1	1	**7**
6	***1***	***1***	***1***	0	0	1	0	***1***	0	0	1	***1***	0	1	1	1	1	**11**
7	***1***	***1***	***1***	0	0	1	1	1	1	0	1	1	0	1	1	1	1	**13**
8	1	***1***	***1***	0	0	0	0	1	0	0	0	0	0	0	0	1	***1***	***6***
9	***1***	***1***	***1***	0	0	0	0	1	1	0	1	0	0	0	0	***1***	***1***	***8***
10	***1***	***1***	***1***	0	0	0	0	1	0	1	0	0	0	0	0	***1***	***1***	***7***
11	***1***	***1***	***1***	0	0	0	0	1	0	0	1	0	0	0	0	***1***	***1***	***7***
12	***1***	***1***	***1***	0	0	0	0	1	0	0	1	1	0	0	0	***1***	***1***	***8***
13	***1***	***1***	***1***	0	0	1	1	***1***	***1***	0	***1***	***1***	1	1	1	***1***	***1***	***14***
14	***1***	***1***	***1***	0	0	0	0	***1***	0	0	1	1	0	1	1	***1***	***1***	***10***
15	***1***	***1***	***1***	0	0	0	0	***1***	0	0	1	1	0	1	1	***1***	***1***	***10***
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	***1***	***1***	***2***
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	**1**
Dependence	13	14	15	1	2	4	3	12	4	2	9	7	2	6	6	13	17	

The bold and italic values show the influencing or driving relationship between Goal i and Goal j through transitivity analysis, i.e. indirect relationship. For example, if Goal 1 influences Goal 2 and Goal 2 influences Goal 3, but Goal 1 does not influences Goal 3 directly. Then, in this situatiuon we use the transitivity analysis and show that Goal 1 also influences the Goal 3 and this relationship is shown by bold and italic values. Only bold values in the last column show the total driving power of the corresponding goal

Table 6 Level partitions of SDGs

Number of iterations	Reachability set	Antecedent set	Interaction set	Level
4	1, 2, 3, 17	1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	1	IV
3	2, 3, 17	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	2	III
2	3, 17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	3	II
12	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	4	4	XII
6	1, 2, 3, 5, 8, 16, 17	4, 5	5	VI
9	1, 2, 3, 6, 8, 11, 12, 14, 15, 16, 17	4, 6, 7, 13	6	IX
10	1, 2, 3, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17	4, 7, 13	7	X
5	1, 2, 3, 8, 16, 17	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15	8	V
7	1, 2, 3, 8, 9, 11, 16, 17	4, 7, 9, 13	9	VII
6	1, 2, 3, 8, 10, 16, 17	4, 10	10	VI
6	1, 2, 3, 8, 11, 16, 17	4, 6, 7, 9, 11, 12, 13, 14, 15	11	VI
7	1, 2, 3, 8, 11, 12, 16, 17	4, 6, 7, 12, 13, 14, 15	12	VII
11	1, 2, 3, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17	4, 13	13	XI
8	1, 2, 3, 8, 11, 12, 14, 15, 16, 17	4, 6, 7, 13,14, 15	14, 15	VIII
8	1, 2, 3, 8, 11, 12, 14, 15, 16, 17	4, 6, 7, 13, 14, 15	14, 15	VIII
2	16, 17	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	16	II
1	17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	17	I

environmental cleanliness and all types of the inequalities prevalent in the society. It is also observed that this Goal 4 lies in the bottom of the hierarchy as prepared using ISM and is, therefore, the most powerful driver. Most people in developing and least developed countries do not have access to rudimentary infrastructure including those related to clean water and sanitation, food, health facility, and quality education. Also, Responsible Consumption and Production (i.e., Goal 12) affects the Life on Land (Goal 15) as well as the Life below Water (Goal 14). Thus, poverty is largely influenced by poor infrastructure, lack of inclusive growth, the absence of decent work and economic growth, lack of sustainable cities, differences between gender and social inequality. In fact, alleviation of poverty would help achieve the goal of Zero Hunger (Goal 2). Thus, it is observed that all the goals are interrelated to each other, as is also depicted in the SSIM.

7 MICMAC analysis

In the MICMAC analysis, there are four clusters, which are explained as follows:

- Autonomous goals: depicting weak driving power and weak dependence.

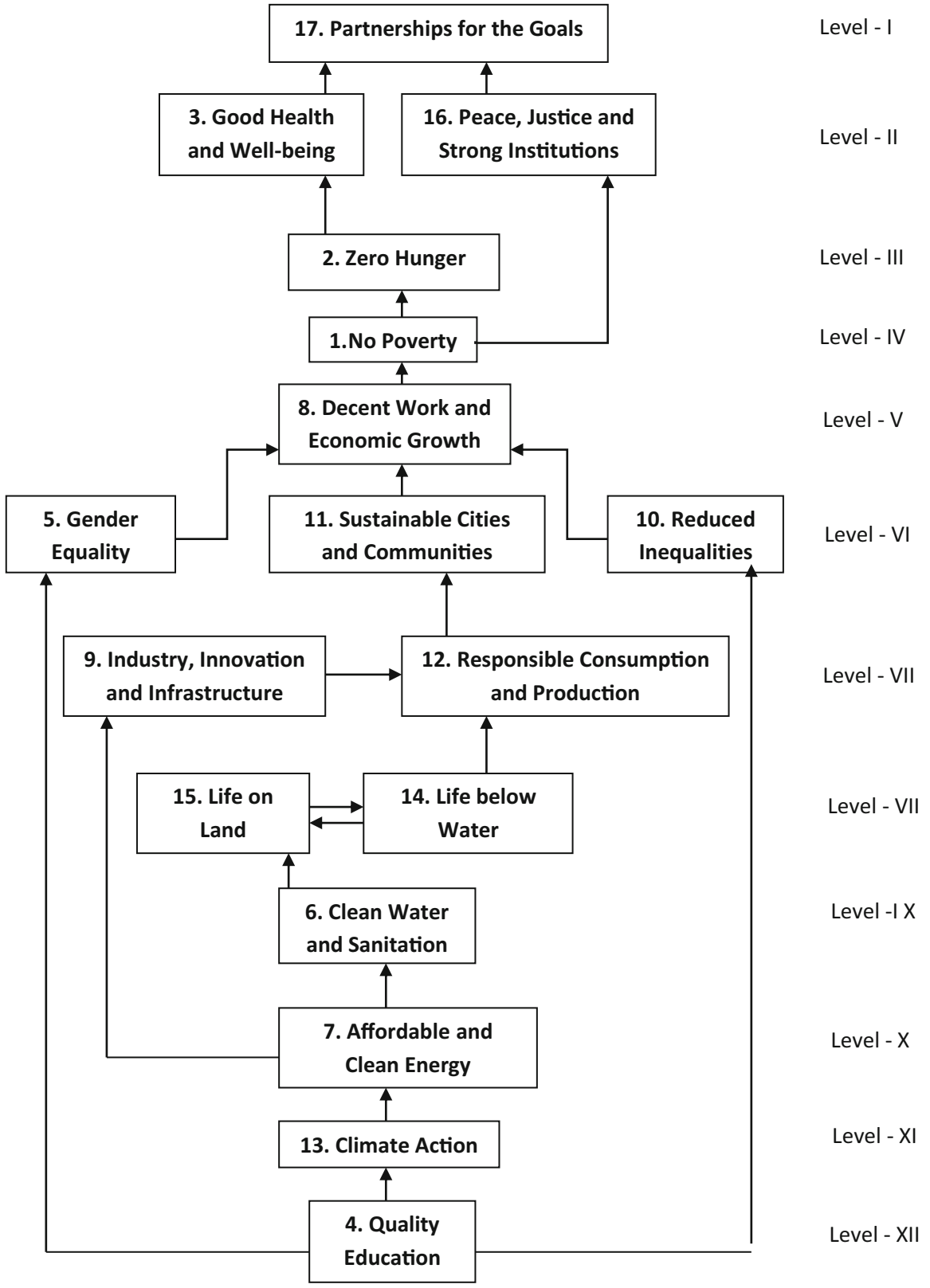


Fig. 1 Hierarchy of sustainable development goals (SDGs)

- Dependent goals: depicting weak driving power but strong dependence.
- Linkage goals: depicting strong driving power and strong dependence.
- Independent goals: depicting strong driving power but weak dependence.

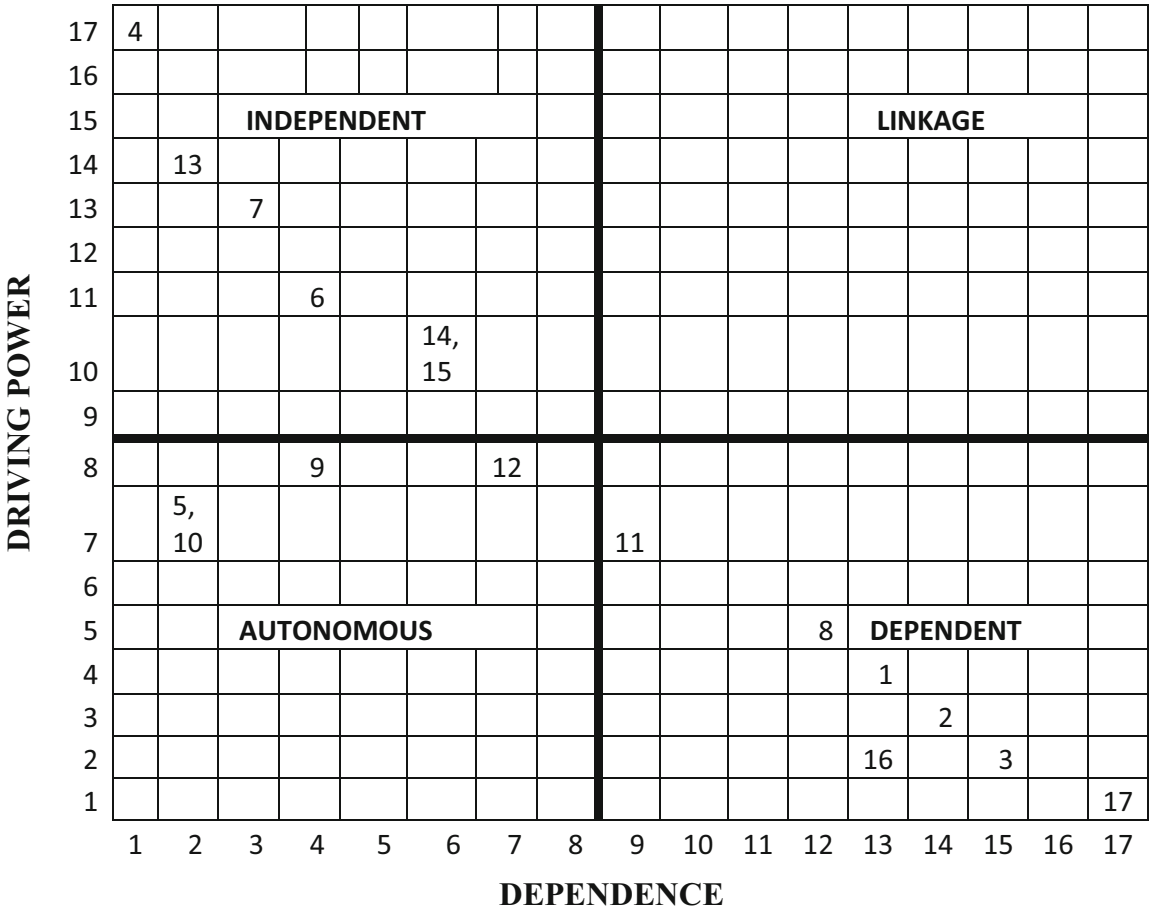


Fig. 2 MICMAC analysis of SDGs

In fact, the goals with a strong driving power are called key goals, and they lie in the category of independent or linkage goals. For an example, it is observed from Table 6 that Goals 14 and 15 have a driving power of 10 and a dependence power of 6. Therefore, in this figure, it is positioned at a place corresponding to a driver power of 10 and a dependence power of 6. Similarly, Goal 17 has the highest dependence power of 17 but the lowest driving power of 1, as shown in Fig. 2.

8 Conclusions

In this study, all the SDGs are presented in the form of a hierarchy, as all goals exhibit a unique mutual inter-relationship. Therefore, prior to implementing a strategic plan to achieve the SDGs, it is necessary for the policymakers in developing and least developed countries to know the level of interaction among the goals and how they influence each other. There should be a sequential relationship among these goals, and it is very important to know the prerequisites before attempting the achievement of a specific goal. There are some goals which cannot be achieved before addressing its prerequisite goal. For instance, if we want to achieve zero hunger, we cannot achieve this goal without alleviating poverty. Similarly, we cannot achieve the goals of gender equality, or reduced inequalities, and climate action without provisions of quality education to the citizens. Quality education, as per our study, is the most rudimentary goal, which instills a motivation among the population to strive for the cause of sustainable development and to save the planet for

generations to come. In this paper, ISM technique has been employed to establish this complex relationship. This technique makes it convenient for policymakers to plan and allocate resources by considering the hierarchy of each SDG as well as their driving and dependence powers.

This paper focuses on the implementation side of the SDGs by respective governments in the global South and intends to provide a working tool for them in their decision-making process. However, the policymakers also need to simultaneously address a number of other constraints viz. financial inclusion, social awareness, political will that could provide auxiliary support while implementing SDGs strategic plans. Kumar et al. (2016) have also highlighted some such pertinent challenges which solicit global attention. These include the following: cost involved in fulfilling some of the SDGs is huge; maintaining peace is essential for development; measuring progress, and ensuring accountability. In our future studies, we intend to employ another suitable modeling technique and analyze some challenges pertinent to SDG implementation and to suggest a perspective that could help policymakers.

Acknowledgement Authors acknowledge the following experts for their opinions shared with us in this research paper: Prof. Anil Kumar Haritas, Department of Environmental Engineering, Delhi Technological University, Delhi. Prof. A. N. Sarkar, Environmental and Social issues expert, Asia-Pacific Institute of Management, New Delhi. Prof. Arun Tripathi, Management Development Institute, Gurugram, Haryana. Dr. K. R. Kaushik, Gujrat State Fertilizers and Chemicals Limited, New Delhi. Mr. Hamid Ahmad, State Consultant for World Bank assisted project on drinking water and sanitation, Govt. of Karnataka, Bengaluru. Prof. Anurika Vaish, Department of Management Studies, Indian Institute of Information Technology, Allahabad. Prof. Imlakh Saikh, Management Development Institute, Gurugram, Haryana. Dr. Harendra Kumar, Amity University Noida, Uttar Pradesh. Mr. M. Absar Alam, Consultant, Asian Development Bank, New Delhi.

References

Akenji, L., & Bengtsson, M. (2014). Making sustainable consumption and production the core of sustainable development goals. *Sustainability, 6*(2), 513–529.

Annan, K. A. (2002). Toward a sustainable future. *Environment, 44,* 10–15.

Berger, G. (2009). Governance for sustainable development: Concepts, principles and challenges. Paper presented at the Eurofound expert meeting: Industrial relations and sustainability, Brussels.

Bouma, J. (2014). Soil science contributions towards sustainable development goals and their implementation: Linking soil functions with ecosystem services. *Journal of Plant Nutrition and Soil Science, 177*(2), 111–120.

Cicin-Sain, B., Necht, R. W., & Fisk, G. W. (1995). Growth in capacity for integrated coastal management since UNCED: An international perspective. *Ocean and Coastal Management, 29*(1–3), 93–123.

Costanza, R., Daly, L., Fioramonti, L., Giovannini, E., Kubiszewski, I., Mortensen, L. F., et al. (2016a). Modelling and measuring sustainable wellbeing in connection with the UN sustainable development goals. *Ecological Economics, 130,* 350–355.

Costanza, R., Fioramonti, L., & Kubiszewski, I. (2016b). The UN sustainable development goals and the dynamics of well-being. *Frontiers in Ecology and the Environment, 14*(2), 59.

Costanza, R., McGlade, J., Lovins, H., & Kubiszewski, I. (2014). An overarching goal for the UN sustainable development goals. *Solutions, 5*(4), 13–16.

Esquivel, V. (2016). Power and the sustainable development goals: A feminist analysis. *Gender & Development, 24*(1), 9–23.

Esquivel, V., & Sweetman, C. (2016). Gender and the sustainable development goals. *Gender & Development, 24*(1), 1–8.

Fredman, S., Kuosmanen, J., & Campbell, M. (2016). Transformative equality: Making the sustainable development goals work for women. *Ethics & International Affairs, 30*(02), 177–187.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., et al. (2013). Sustainable development goals for people and planet. *Nature: International Weekly Journal of Science, 495*, 305–307.

Griggs, D., Stafford Smith, M., Rockström, J., Öhman, M. C., Gaffney, O., Glaser, G., et al. (2014). An integrated framework for sustainable development goals. *Ecology and Society, 19*(4), 49. doi:10.5751/ES-07082-190449.

Haines, A., & Cassels, A. (2004). Can the millennium development goals be attained? *British Medical Journal, 329,* 397–994.

Joshi, D. K., Hughes, B. B., & Sisk, T. D. (2015). Improving governance for the post-2015 sustainable development goals: Scenario forecasting the next 50 years. *World Development, 70,* 286–302.

Kamepalli, L. B., & Pattanayak, S. K. (2015). From millennium to sustainable development goals and need for institutional restructuring. *Current Science, 108*(6), 1043–1044.

Klapper, L., El-Zoghbi, M., & Hess, J. (2016). *Achieving the sustainable development goals: The role of financial inclusion* (pp. 1–15). Working Paper. Retrieved from https://www.unsgsa.org/files/3114/6489/8605/sdgs_paper_final_003.pdf.

Kumar, S., Kumar, N., & Vivekadhish, S. (2016). Millennium development goals (MDGS) to sustainable development goals (SDGS): Addressing unfinished agenda and strengthening sustainable development and partnership. *Indian Journal Of Community Medicine: Official Publication of Indian Association of Preventive & Social Medicine, 41*(1), 1–4.

Lange, P., Driessen, P. P., Sauer, A., Bornemann, B., & Burger, P. (2013). Governing towards sustainability—Conceptualizing modes of governance. *Journal of Environmental Policy & Planning, 15*(3), 403–425.

Malik, O. A., Hsu, A., Johnson, L. A., & de Sherbinin, A. (2015). A global indicator of wastewater treatment to inform the sustainable development goals (SDGs). *Environmental Science & Policy, 48,* 172–185.

Mandal, A., & Deshmukh, S. G. (1994). Vendor selection using interpretive structural modeling (ISM). *International Journal of Operations & Production Management, 14*(6), 52–59.

MDG Report. (2015). *Foreword by UN Secretary General Ban Ki-Moon.* Available http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf.

Niestroy, I. (2014). Governance for sustainable development: How to support the implementation of SDGs? Chapter 10 of ASEF outlook report 2014–2015 facts and perspectives (pp. 151–165). Retrieved from http://www.asef.org/images/docs/ASEM%20Outlook%202015%20Vol%20II.pdf#page=156.

Norheim, O. F., Jha, P., Admasu, K., Godal, T., Hum, R. J., Kruk, M. E., et al. (2015). Avoiding 40% of the premature deaths in each country, 2010–30: Review of national mortality trends to help quantify the UN sustainable development goal for health. *The Lancet, 385*(9964), 239–252.

Norström, A. V., Dannenberg, A., McCarney, G., Milkoreit, M., Diekert, F., Engström, G., Fishman, R., Gars, J., Kyriakopoolou, E., Manoussi, V., & Meng, K. (2014). Three necessary conditions for establishing effective sustainable development goals in the anthropocene. *Ecology and Society, 19*(3), 8.

Osborn, D., Cutter, A., & Ullah, F. (2015). Universal sustainable development goals: Understanding the transformational challenge for developed countries. Report of a Study by Stakeholder Forum (pp. 5–20).

Osborn, D., & Ullah, F., (2014). Sustainable development goals. A Universal Agenda that should shape strategies, priorities and action in both developed and developing countries, and in middle income countries. How to build on the Draft Goals prepared by the UN's open working group—A challenge—And an opportunity—For Europe. Paper by Stakeholder Forum for ESDN conference. Rome 6–7 November 2014.

Pablos-Mendez, A., Cavanaugh, K., & Ly, C. (2016). The new era of health goals: Universal health coverage as a pathway to the sustainable development goals. *Health Systems & Reform, 2*(1), 15–17.

Pinter, L., Almassy, D., Antonio, E., Niestroy, I., Olsen, S., & Pulawska, G. (2013). Sustainable development goals and indicators for a small planet. Part I: Methodology and goal framework, Interim Report (pp. 1–84). Retrieved from http://sdg.earthsystemgovernance.org/sdg/sites/default/files/files/publications/Sustainable%20Development%20Goals%20and%20Indicators%20for%20a%20Small%20Planet-Interim%20Report.pdf.

Pisano, U., Lange, L., Berger, G., & Hametner, M. (2015). The sustainable development goals (SDGs) and their impact on the European SD governance framework. ESDN Quaterly report (pp. 1–65). Retrieved from http://www.sd-network.eu/quarterly%20reports/report%20files/pdf/2015-January-The_Sustainable_Development_Goals_(SDGs)_and_their_impact_on_the_European_SD_governance_framework.pdf.

Sachs, J. D. (2012). From millennium development goals to sustainable development goals. *The Lancet, 379*(9832), 2206–2211.

Steurer, R. (2009). Sustainable development as governance reform agenda; an aggregation of distinguishes challenges for policy-making. Institute of forest, environment and natural resource poliy. Discussion paper (pp. 1–18). Retrieved from http://www.wiso.boku.ac.at/fileadmin/data/H03000/H73000/H73200/InFER_Discussion_Papers/InFER_DP_09_1_SD_as_governance_reform_agenda_final.pdf.

Stokstad, E. (2015). Sustainable goals from UN under fire. *Science, 347*(6223), 702–703.

Szabo, S., Renaud, F. G., Hossain, M. S., Sebesvári, Z., Matthews, Z., Foufoula-Georgiou, E., et al. (2015). Sustainable development goals offer new opportunities for tropical delta regions. *Environment: Science and Policy for Sustainable Development, 57*(4), 16–23.

Thornicroft, G., & Patel, V. (2014). Including mental health among the new sustainable development goals. *BMJ, 349*(7972), 5.

UNGA. (2015). *Transforming our world: The 2030 Agenda for sustainable development*. Resolution adopted by the general assembly on 25 September 2015, seventieth session Agenda items 15 and 116. Available: A/RES/70/1, http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E.

United Nations (UN). (2012). *The future we want: Outcome document adopted at Rio+20*. United Nations, New York, New York, USA. [online] URL http://www.un.org/futurewewant.

Votruba, N., Eaton, J., Prince, M., & Thornicroft, G. (2014). The importance of global mental health for the sustainable development goals. *Journal of Mental Health, 23*(6), 283–286.

Waage, J., Banerji, R., Campbell, O., Chirwa, E., Collender, G., Dieltiens, V., et al. (2010). The Millennium Development Goals: A cross-sectoral analysis and principles for goal setting after 2015. *The Lancet, 376*(9745), 991–1023.

WCED. (1987). *World commission on environment and development. Our common future* (p. 9). Oxford: Oxford University Press.

Sensors and Actuators B 253 (2017) 115–123

Contents lists available at ScienceDirect

Sensors and Actuators B: Chemical

journal homepage: www.elsevier.com/locate/snb

Research paper

Development of a paper-based electrochemical immunosensor using an antibody-single walled carbon nanotubes bio-conjugate modified electrode for label-free detection of foodborne pathogens

Jyoti Bhardwaj[a], Sivaranjani Devarakonda[b], Suveen Kumar[b], Jaesung Jang[a,b,*]

[a] Department of Biomedical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea
[b] School of Mechanical, Aerospace and Nuclear Engineering, UNIST, Ulsan 44919, Republic of Korea

ARTICLE INFO

Article history:
Received 4 February 2017
Received in revised form 22 May 2017
Accepted 15 June 2017
Available online 21 June 2017

Keywords:
Antibody-carbon nanotube bio-conjugate
Staphylococcus aureus
Paper substrate
Electrochemical immunosensor
Label-free detection

ABSTRACT

The need for low-cost, sensitive, and reliable sensors for the detection of whole bacterial cells in food samples without pre-treatment has been increasing. Outbreaks of foodborne diseases can be severe, especially in developing countries; however, most bio-detection tools are unaffordable. Here, we have developed a rapid and low-cost paper-based electrochemical immunosensor for label-free detection of *Staphylococcus aureus*, using antibody (Ab)-single walled carbon nanotube (SWCNT) bio-conjugates. Anti-*S. aureus* antibodies were covalently attached onto the SWCNTs, using the *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride/*N*-hydroxysuccinimide coupling reagent. These Ab-SWCNT bio-conjugates were then immobilized on the working electrode, and the presence of *S. aureus* was detected by analyzing the change in peak current following antigen-antibody complex formation. Differential pulse voltammetry was performed with a bacterial concentration ranging from 10 to 10^7 colony forming units (CFU) mL^{-1}. A selectivity assay using *Escherichia coli* B, *Bacillus subtilis*, and *S. epidermidis* (to examine cross-reactivity) showed that the sensor was specific to *S. aureus*. Moreover, this immunosensor showed a good linear relationship ($R^2 = 0.976$) between the increase in peak current and logarithmic *S. aureus* concentration, with a rapid detection time (30 min) and a limit of detection of 13 CFU mL^{-1} in spiked milk samples. This low-cost immunosensor can be used for rapid detection of pathogens in actual food samples with high sensitivity and specificity.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

There has been a growing demand for the development of sensitive and low-cost sensors for the detection of pathogenic foodborne microorganisms in a wide variety of applications such as those in the food industry [1]. *Campylobacter spp., Salmonella spp.*, *Escherichia coli* O157:H7, *Staphylococcus aureus*, and *Listeria monocytogenes* are considered the most infectious foodborne bacteria [2]. Of these, *S. aureus* is one of the major foodborne bacteria responsible for a broad range of diseases; *S. aureus*-contaminated meat and milk-based products are the second major cause, after *Salmonella*, of food poisoning outbreaks in the UK and France [3]. *S. aureus* has been shown to express multiple virulence factors in the form of toxins in several food samples [4]. The infectious dose for *S. aureus* is reported to be 10^3–10^8 cells for humans [5]; the hazardous level of *S. aureus*-released enterotoxins is produced by 10^8 cells per kg of food [6] and 20 ng of enterotoxin present in food can cause mild intoxication [7]. The acceptable limit of *S. aureus* is up to 100 colony forming units (CFU) per gram [8].

The first and critical step in controlling pathogenic foodborne bacteria is to effectively and rapidly detect the number of bacteria in food samples without the need for pre-treatment [9]. The available conventional and advanced techniques are sufficiently sensitive; however, they are usually time-consuming, expensive, and require trained personnel and multiple pre-treatment steps (such as extraction, isolation, enrichment, and counting) [10,11]. Therefore, there is an increasing need for low-cost, sensitive, and reliable sensors able to detect whole bacterial cells in food samples. Outbreaks of foodborne diseases can be severe, especially in developing countries; however, most sensitive bio-detection tools are unaffordable.

In this regard, paper-based biosensors are a promising alternative because of the advantageous characteristics of paper

* Corresponding author at: School of Mechanical, Aerospace and Nuclear Engineering, UNIST, Ulsan, 44919, Republic of Korea.
E-mail address: jjang@unist.ac.kr (J. Jang).

http://dx.doi.org/10.1016/j.snb.2017.06.108
0925-4005/© 2017 Elsevier B.V. All rights reserved.

substrates, such as flexibility, lightness, and a fibrous and porous structure [12]. Moreover, paper is inexpensive, biocompatible, biodegradable, and easily patterned using printing technology [13]. These properties enable the creation of a low-cost and portable sensing device for several applications. Several colorimetric paper-based sensors are commercially available, such as home pregnancy test kits; however, these sensors usually provide qualitative detection [14]. Hence, much research effort has been directed towards developing paper-based biosensors that can also offer quantitative information [15]. Electrochemical detection techniques are well suited for use in conjunction with paper-based biosensors because they are cost-effective, do not require expensive instrumentation, and can achieve high sensitivity [16].

Many studies have been conducted for quantification of *S. aureus*, using various platforms [17]. A surface plasmon resonance-based biosensor was developed for the detection of *S. aureus*, using a lytic phage as a specific and selective probe [18]. Tan et al. [19] developed a microfluidic immunosensor chip for the detection of *S. aureus*. Braiek et al. [20] used a self-assembled monolayer of alkane thiols on a gold surface and performed electrochemical impedance spectroscopy for immunodetection of *S. aureus* with a detection limit of 10 CFU mL^{-1}. Gomez et al. [21] reported an enzyme-labelled amperometric immunosensor for the detection of *S. aureus* with a detection limit of 100 CFU mL^{-1}. Abbaspour et al. [22] developed a dual aptamer based electrochemical immunosensor for *S. aureus*, obtaining a detection limit of 1 CFU mL^{-1}. However, most paper-based sensors used to detect infectious bacteria involved colorimetric and fluorometric detection methods, which are either qualitative or semi-quantitative. In fact, paper-based electrochemical immunosensors for bacterial detection have received little attention (Table 1).

In this study, we demonstrate a low-cost paper-based electrochemical biosensor modified with antibody (Ab)-single walled carbon nanotube (SWCNT) bio-conjugates for label-free detection of *S. aureus*. Several techniques have been developed for constructing paper-based immunosensors by different hydrophobic-hydrophilic patterning methods, including photolithography [23], inkjet printing [24], stamping, cutting [25], screen-printing [26], and wax printing [27]. Wax printing, the technique applied to create hydrophobic barriers in the present study, is a rapid, low-cost, and simple mass production fabrication method [28]. However, the hydrophobic patterns deposited on paper substrates can be damaged by excessive use of solvents (or washing steps) during the immobilization and immunoassay steps [29]. To reduce the number of immunoassay processing steps, thereby helping to maintain the hydrophobic patterns and increasing the robustness of the paper-based sensor, the Ab-SWCNT bio-conjugates were prepared using covalent interactions between the SWCNTs and antibodies specific to *S. aureus* via a diimide-activated amidation process [30]. Ab-SWCNT bio-conjugates have been developed previously for selective targeting of tumor cells, using non-covalent interactions via physical adsorption [31], covalent bonding [32], or indirectly via streptavidin-biotin interaction [33]. Antibody-functionalized SWCNTs were also used to demonstrate electrochemical measurements for *Salmonella* with a glassy carbon electrode; however, the electrochemical assay was not fully characterized [34].

The present immunosensor was fabricated via the deposition of carbon working and counter electrodes and an Ag/AgCl reference electrode on paper using stencil printing. Carbon is an attractive material for electrodes because of its inert nature, low cost, simple fabrication, low background current, and easy surface renewability for electron transfer [35]. The Ab-SWCNT bio-conjugate was then deposited on the working electrode, and sensor performance (including sensitivity, selectivity, and stability) was measured using

Table 1
A list of paper-based electrochemical immunosensors for bacterial detection and previous electrochemical immunosensors specific for *S. aureus* detection.

Type of Sensor	Substrate	Immobilization Scheme	Ag-Ab interaction	Signal Amplification	Bacteria	Limit of Detection (CFU mL^{-1})	Dynamic Range (CFU mL^{-1})	References
Impedance	Gold electrode	SAM formation	Direct EIISA (Ag-Ab)	Label free	*S. aureus*	10	$10–10^6$	Braiek et al. (2012)
Impedance.	Alumina nonporous membrane	SAM formation	Direct EIISA (Ag-Ab)	Label free	*S. aureus*	10^2	$10^2–10^7$	Tan et al. (2011)
Impedance	Gold electrode	Electro grafting of p-amino phenyl (p-AP) film on gold electrode	Direct EIISA (Ag-Ab)	Label free	*S. aureus*	10	$10–10^7$	Chrouda et al. (2013)
Amperometric	Gold electrode	SAM formation using DTSP	Competitive ELISA	Enzyme labelled antibody	*S. aureus*	10^2	$1.3 \times 10^3–7.6 \times 10^4$	Gomez et al. (2008)
Differential pulse voltammetry	Carbon working electrode	–	Dual Aptamer based Sandwich ELISA	Silver nanoparticle labelled aptamer	*S. aureus*	1	$10–10^7$	Abbaspour et al. (2015)
Impedance	Graphene paper electrode	Modified with gold nanoparticles	Sandwich ELISA (Biotin-Streptavidin interaction)	Label free	*E.coli* O157:H7	10^2	$10^2–10^7$	Wang et al. (2013)
Impedance	Graphene paper electrode	Modified with platinum nano- cauliflower	RNA Aptamer	Label free	*E.coli* O157:H7	4	$1–10^6$	Burrs et al. (2016)
Differential pulse voltammetry	Paper based carbon paste electrode	Modified with SWCNT-Ab conjugates	Direct EIISA (Ag-Ab)	Label free	*S. aureus*	15	$10–10^7$	Present work

phosphate-buffered saline (PBS) and milk contaminated with *S. aureus*.

2. Materials and methods

2.1. Materials

SWCNTs (90% SWCNTs, diameter: 1–2 nm, length: 5–30 μm, COOH content: 2.75 wt%) were obtained from MK Impex corp. (Canada). *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide hydrochloride (EDC), *N*-hydroxysuccinimide (NHS), Whatman chromatography paper 3 mm (WHA3030861), bovine serum albumin (BSA; A2153), and Tween-20 were obtained from Sigma-Aldrich (USA). Dimethylformamide (DMF; 98%, D1021) and PBS (1×, pH 7.4) were purchased from Biosesang Inc. (South Korea). Conductive carbon paste (DC-21, sheet resistance: 20–25 ohm/square at 20 μm thickness) was obtained from Dozen tech (South Korea) and Ag/AgCl ink (RD66) was purchased from BAS Inc. (Tokyo, Japan). Rabbit anti-*Staphylococcus aureus* antibody (polyclonal antibody; ab20920) was purchased from Abcam (UK). *Staphylococcus aureus* reference strain (ATCC 25923), *Escherichia coli* B (ATCC 11303), *Bacillus subtilis* (ATCC 21332), and *Staphylococcus epidermidis* (ATCC 12228) were obtained from the Korean Collection for Type Cultures (South Korea).

2.2. Bacterial cultivation

Bacterial strains (*S. aureus*, *E. coli* B, *B. subtilis*, and *S. epidermidis*) were grown on Luria-Bertani agar (LB, Difco) plates for 24 h at 37 °C. A single colony was isolated and inoculated into a 50-mL conical tube containing LB broth, and was grown overnight at 37 °C in a shaking incubator. One milliliter of the bacterial culture was then used to inoculate fresh LB broth in another tube, which was then incubated overnight with shaking at 37 °C. Bacterial growth was determined by measuring optical density (OD) at 600 nm (OD_{600}). When the OD_{600} of the bacterial strains reached 0.6 (corresponding to 4.2×10^8 cells mL^{-1}), 1 mL of the bacterial culture (per strain) were used to inoculate fresh LB broth tubes, which were incubated at 37 °C with shaking for 6 h prior to each experiment. Twenty milliliters of these bacterial cultures were centrifuged for 10 min at 4000 rpm (3220×*g*) and the pellets were washed twice with PBS. Finally, the bacterial pellets were re-suspended in 20 mL of PBS and a serial dilution was prepared in PBS.

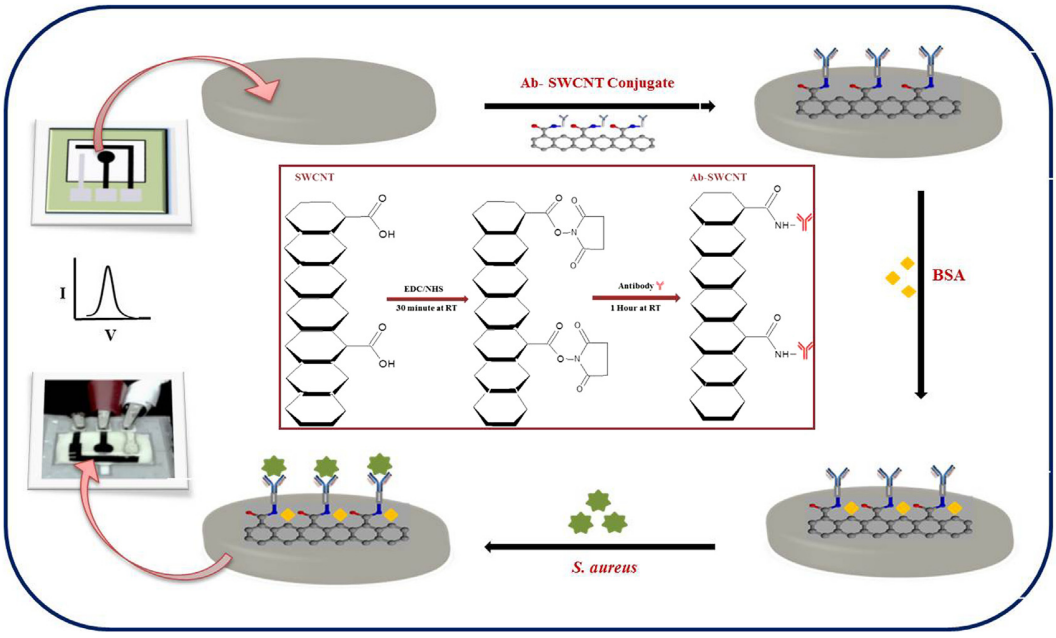
2.3. Sensor fabrication

The design of the present paper sensor was created using AutoCAD software and printed on Whatman paper, using a solid wax printer (ColorQube 8570, Xerox). Colored areas covered with wax were hydrophobic, whereas colorless areas yielded hydrophilic features on the paper. The printed paper was then baked on a hot plate at 140 °C for 2 min, allowing the wax to diffuse through the porous paper and creating hydrophobic barriers. The patterned paper was then removed from the hot plate and cooled to room temperature.

The three electrochemical assay electrodes were created by stencil printing. Carbon paste was used for constructing the working (WE) and counter electrodes (CE), and silver/silver chloride (Ag/AgCl) ink was used for the reference electrode (RE) and contact pads on the patterned paper (Scheme 1). The stencils for the electrode patterns were formed using AutoCAD software followed by cutting with a laser cutter (Versa Laser, USA). The 100-μm thick adhesive polyester stencil mask was placed on the paper with the electrode sensing areas on the hydrophilic section and the contact pads on the hydrophobic section, and the carbon paste was then spread uniformly. Next, the stencil mask was peeled off after air-drying at room temperature for 10 min, and the paper was baked in an oven at 60 °C for 30 min. Ag/AgCl ink was also applied to create the RE and dried in an oven at 50 °C for 20 min. The thickness of the electrodes was approximately 100 μm, and the areas of the WE, CE, and RE were 19.94 mm^2, 42.08 mm^2, and 15.68 mm^2, respectively.

2.4. Conjugation of the Antibody to the Carbon Nanotube

Simple and rapid fabrication of the immunosensor on the paper substrate was achieved by conjugating the SWCNTs with antibodies via covalent bonding (Scheme 1) according to the method pre-



Scheme 1. Schematic of the Ab-SWCNTs bio-conjugate: attachment of antibodies to SWCNTs using EDC/NHS coupling reagent, electrochemical biosensor fabrication, and assay procedure.

viously described by Jiang et al. [30]. Briefly, the SWCNTs were dispersed in DMF (2 mg mL^{-1}) by sonication for 2 h in a water bath, centrifuged at 5000 rpm for 1 h, and the supernatant was then collected [36–38]. EDC/NHS coupling reagents were used for covalent bonding of the primary amines of antibodies with the —COOH groups of the SWCNTs. First, the SWCNTs were incubated for 30 min with 1 mL of a 400 mM EDC and 100 mM NHS mixture (1:1 vol ratio) to activate the carboxylic groups on the SWCNTs and to produce a highly reactive and stable amine-reactive NHS ester. Subsequent to washing with PBS, the activated SWCNTs were collected by centrifugation at 5000 rpm for 10 min and incubated with the anti-*S. aureus* antibody (5 μg mL^{-1}) for 2 h at room temperature followed by centrifugation at 15000 rpm at 4 °C for 10 min. The supernatant was discarded and 1 mL of 0.05% (v/v) Tween-20 in PBS was added to the conjugates. These conjugates were vortexed and stored at 4 °C. The stability and covalent binding of these Ab-SWCNT conjugates were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The presence of dissociated protein was detected by staining with Coomassie blue [39]. A mixture of the anti-*S. aureus* antibody and non-activated carboxylated SWCNTs (non-covalent bonding or without EDC/NHS activation) was used as a control.

2.5. Immobilization

Two microliters of the Ab-SWCNT conjugate was dropped on the carbon paste electrode (CPE) and incubated for 30 min at 37 °C. Excess and unbound antibodies were removed by washing with PBS, and the CPE was dried under N_2 gas. The Ab-SWCNT modified CPE was then treated with BSA (1 mg mL^{-1}) for 30 min at room temperature to prevent nonspecific binding. After rinsing with PBS, the CPE was dried under an N_2 gas flow and incubated with various concentrations of *S. aureus* ranging from 10 to 10^7 CFU mL^{-1} diluted in PBS for 30 min at 37 °C. Finally, the electrode surface was washed with PBS and dried under N_2 gas. For selectivity studies, the electrode was incubated with *E. coli* B, *B. subtilis*, and *S. epidermidis* following the same protocol as that for *S. aureus*.

2.6. Electrochemical characterization

Differential pulse voltammetry (DPV) and cyclic voltammetry (CV) were performed using Autolab, a potentiostat/galvanostat, and NOVA 1.10 Software (Metrohms, Netherlands). All electrochemical measurements were carried out at room temperature (23 ± 2 °C) using 10 mM PBS (1×, pH 7.4) containing 1 mM $[Fe(CN)_6]^{-3/-4}$ and 0.1 M KCl. DPV and CV were performed for the electrochemical characterization of CPE, SWCNT/CPE, Ab/SWCNT/CPE, and Ab-SWCNT conjugates/CPE with and without various concentrations of *S. aureus*. All calculations and analyses were carried out with Origin pro software.

2.7. Fourier-Transform Infrared Spectrometry (FTIR) and Scanning Electron Microscopy (SEM)

FTIR (VARIAN 670/620, CA, USA) was performed for the analysis of functional groups on the SWCNTs and bond formation after binding of antibodies (Ab-SWCNT conjugate); the spectra were acquired in the absorbance mode. Field emission SEM (Hitachi S-4800, Japan) was used to confirm the change in paper structure subsequent to wax printing. Structural and morphological assessment of the electrode surface areas was conducted after deposition of SWCNT/CPE, Ab-SWCNT conjugate/CPE, and *S. aureus*/Ab-SWCNT conjugate/CPE. Samples were dried overnight and platinum was sputter-coated on the samples at 20 mA for 60 s to increase the optical resolution.

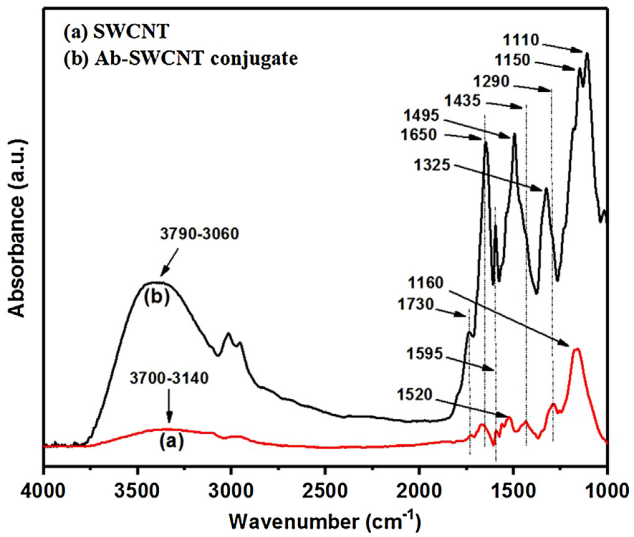


Fig. 1. FT-IR spectra of SWCNTs and Ab-SWCNT bio-conjugates.

3. Results and discussion

3.1. Physiochemical characterization

The hydrophobic and hydrophilic areas were defined on Whatman paper by wax printing, and antibodies specific to *S. aureus* were covalently linked to the SWCNTs via amide bond formation using the crosslinking reagent EDC/NHS. FT-IR studies were carried out to investigate the presence of functional groups on the SWCNTs and Ab-SWCNT conjugates; gel electrophoresis was conducted to confirm the covalent bonding of antibodies on the SWCNTs. In addition, SEM was conducted to verify the change in morphology of the paper after wax printing and functionalization of the electrode surface.

Fig. 1 shows the FT-IR spectra of the (a) SWCNT and (b) Ab-SWCNT conjugate. In both spectra (a) and (b), the peaks obtained at 1290 and 1730 cm^{-1} are due to the —C—O group and —C=O group present in the carboxylic acid of the functionalized SWCNTs, respectively [40]. Furthermore, the peaks obtained in both spectra at 1435 cm^{-1} represent a C—C bond and the peaks at 1595 and 1650 cm^{-1} signify a C=C bond present in the aromatic structure of the SWCNTs [40]. In spectrum (a), the sharp peak at 1160 cm^{-1} indicates the presence of a carboxylic acid C—O group and the broad peak obtained between 3700 and 3140 cm^{-1} indicates the presence of an —OH group in the carboxylic acid [40]. In spectrum (b), the additional peak obtained near 1150 cm^{-1} is due to the formation of a strong amide bond [40] (C—N stretching) between the carboxylic acid present on the surface of the SWCNT and the $-NH_2$ groups present on anti-*S. aureus* antibody, which confirms the successful binding of the antibody to the SWCNTs. Furthermore, the peaks obtained at 1110 and 1325 cm^{-1} may be ascribed to a carboxylic acid —C—O bond present on the Fc region and 1495 cm^{-1} may be due to —C—C bonds present in the anti-*S. aureus* biomolecule. The broad and prolonged peak between 3790 and 3060 cm^{-1} may be due to the cumulative effect of free carboxylic acid —OH groups, $-NH_2$ groups (stretching vibration) of the anti-*S. aureus* molecules, and amide bond formation.

To investigate the covalent bonding of the conjugates, the Ab-SWCNT conjugates were dispersed in sample loading buffer and run under non-reducing conditions. Free antibody (anti-*S. aureus*, 175 kD) and a mixture of antibody (anti-*S. aureus*) and non-activated —COOH SWCNTs were used as the reference and control, respectively (Fig. S1: lanes 1 and 3). As shown in Fig. S1, no antibod-

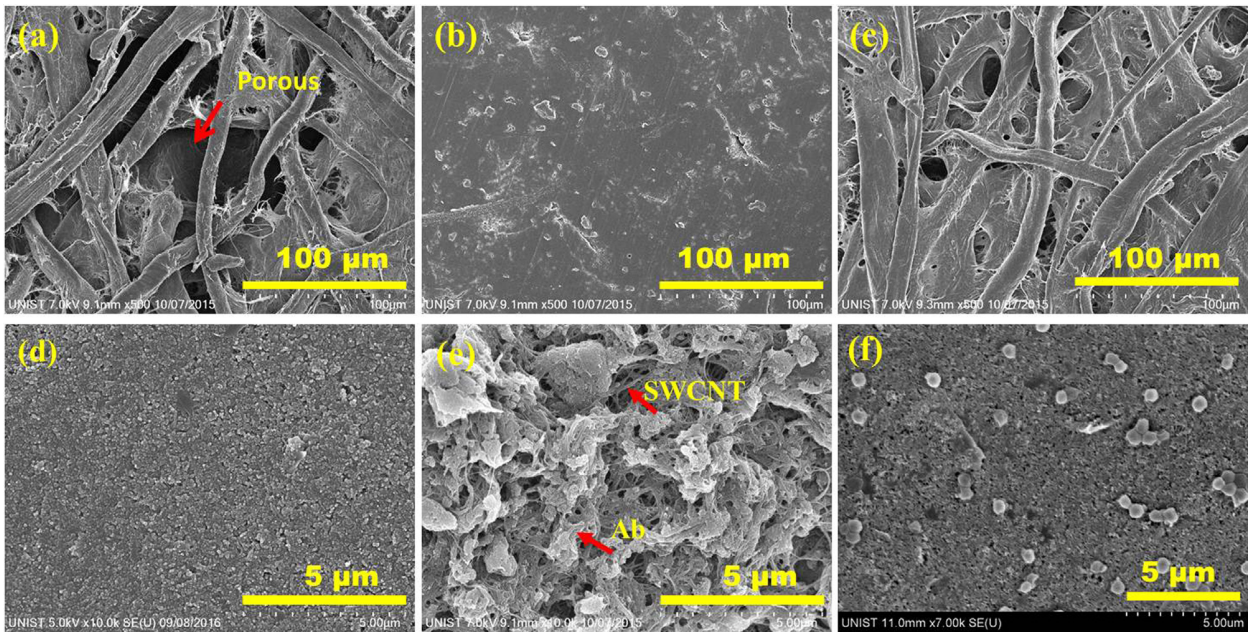


Fig. 2. SEM images for the change in surface morphology following patterning and immobilization. (a) Bare paper. (b) Wax printed paper prior to baking and (c) post baking. (d) Carbon paste electrode (CPE). (e) Ab-SWCNT conjugate/CPE. (f) *S. aureus*/Ab-SWCNT conjugate/CPE.

ies from the Ab-SWCNT conjugates (lane 2) entered the gel, which confirms that the antibodies were covalently attached to the SWCNTs. In contrast, a band corresponding to the antibody was detected in the control sample (lane 3), signifying that the antibodies were non-covalently attached.

The morphology of the functionalized paper was further characterized by SEM. Fig. 2 shows the SEM images of bare Whatman paper (a), wax-printed paper prior to baking (b) and post baking (c), CPE (d), Ab-SWCNT conjugate/CPE (e), and *S. aureus*/Ab-SWCNT conjugate/CPE (f). Paper alone was rather rough and porous (Fig. 2a); however, the surface became smoother following wax printing (Fig. 2b). Subsequent to baking (Fig. 2c), the paper became relatively less porous. A significant change in morphology was apparent following the deposition of the Ab-SWCNT conjugate on the CPE (Fig. 2e); a fibrous SWCNT network was clearly visible beneath the antibodies. The fibrous and globular surface shows the successful immobilization of the Ab-SWCNT conjugates on the CPE. Fig. 2f shows an SEM image of several *S. aureus* bacteria captured by antibodies. The average size of a single *S. aureus* has been reported to be 0.6–1.5 µm [41].

3.2. Characterization of Functionalized Electrodes

Electrochemical characterization was conducted by DPV and CV using $K_3[Fe(CN)_6]$ as an electron transfer mediator. Fig. 3a shows the DPV curves for CPE, SWCNTs/CPE, Ab/SWCNT/CPE, and Ab-SWCNT conjugate/CPE. As demonstrated by the DPV curves, the peak current increased from 12.5 µA (CPE) to 15.8 µA following the deposition of SWCNTs onto the CPE, indicating increased electron transfer between the solution and the SWCNT-modified carbon electrode. SWCNTs exhibit high electrical conductivity and they are used in electrochemical biosensing applications because of their efficient electron transfer during redox reactions [42,43]. Subsequent to the immobilization of antibodies on the SWCNT/CPE, the peak current decreased to 10.4 µA because of the insulating behavior of the antibody that blocked electron transfer. In contrast, the peak current measured for the Ab-SWCNT conjugate/CPE was 14.6 µA; this increased peak current could be due to the defect

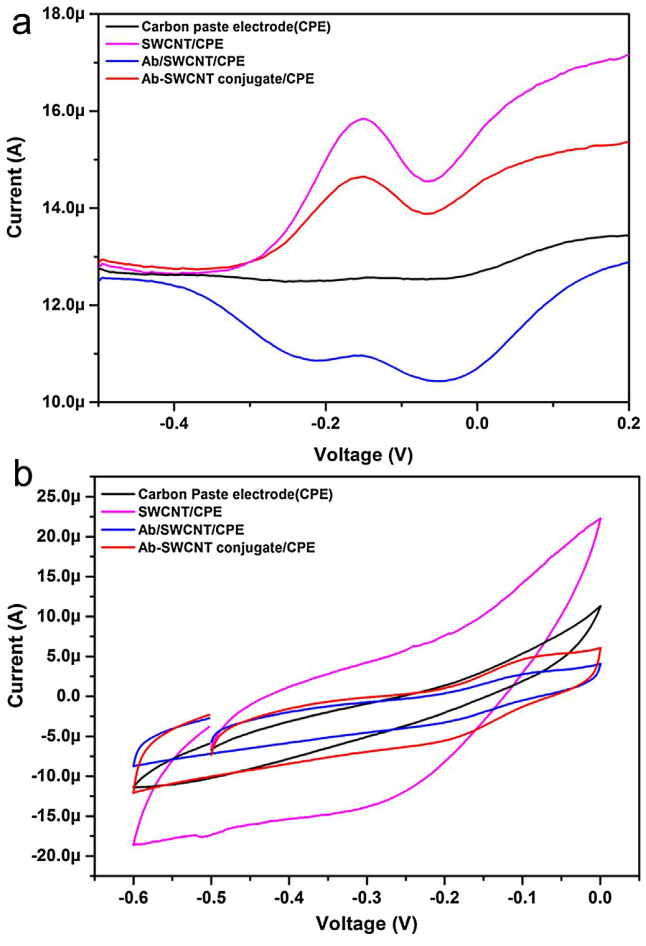


Fig. 3. Differential Pulse Voltammograms (a) and Cyclic voltammograms (b) in the presence of 1 mM $[Fe(CN)_6]^{-3/-4}$ as the redox medium for characterization of CPE, SWCNT/CPE, Ab/SWCNT/CPE, and Ab-SWCNT conjugate/CPE.

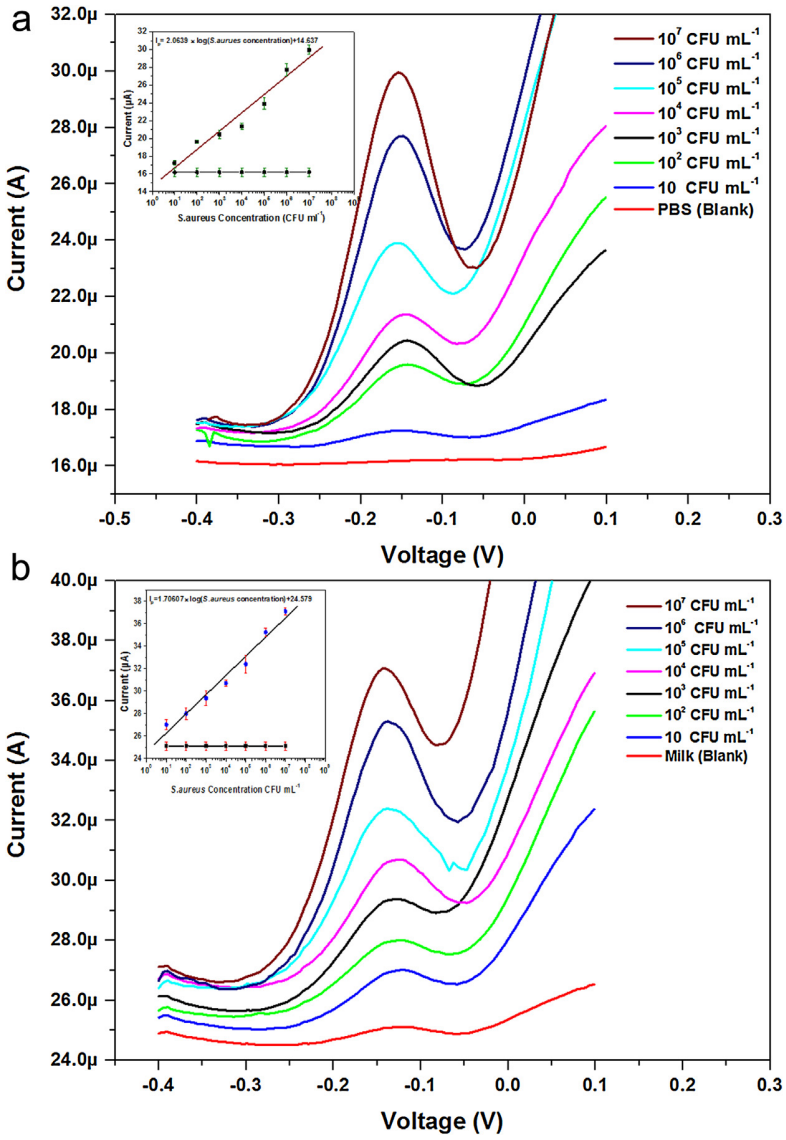


Fig. 4. Sensitivity of the immunosensors. DPVs of varying concentration of *S. aureus* from 10 to 10^7 CFU mL^{-1} in PBS (1×, pH 7.4) containing 1 mM $[Fe(CN)_6]^{-3/-4}$ and 0.1 M KCL (a) in PBS and (b) Milk samples. The error bars represent the standard deviation of three independent measurements.

in antibody monolayer coverage on the SWCNTs, which could enhance the electron transfer between $[Fe(CN)_6]^{-3/-4}$ ions and the electrode surface [44]. The CV measurements were in good agreement with the DPV measurements; the peak current increased following SWCNT deposition and decreased subsequent to the deposition of Ab-SWCNT conjugates (Fig. 3b). DPV was also performed to investigate the response time of the immunosensor for *S. aureus* and the Ab-SWCNT conjugate/CPE modified electrode by varying incubation time up to 60 min (Fig. S2). This result showed that an incubation time of approximately 30 min was required for complete interaction between *S. aureus* and the Ab-SWCNT conjugate/CPE. However, after 30 min, the peak current became nearly saturated.

3.3. *Differential pulse voltammograms for* S. aureus *detection*

DPV electrochemical characterization was conducted for *S. aureus* detection. Five microliters of a *S. aureus* bacterial culture diluted in PBS (1×, pH 7.4) was applied onto the Ab-SWCNT conjugate modified CPE (the bacterial concentration varied from 10 to 10^7 CFU mL^{-1}), and corresponding voltammograms were recorded after 30 min incubation with *S. aureus* at 37 °C. The blank response was obtained when the immunosensor was incubated in PBS without *S. aureus* (Fig. 4a). As shown in Fig. 4a, the peak current increased gradually as the bacterial concentration increased. The peak current changes were due to the formation of an antigen–antibody complex at the electrode surface [45]. A similar trend of increased current has been observed for other biological analytes, such as prostate specific antigen [44], carcinoembryonic antigen [46], and influenza virus H1N1 [47], in several label-free electrochemical detection studies, although a clear mechanism has yet to be established.

There are three possible mechanisms for the label-free electrochemical detection in this study. The increase in current might be caused by defects in monolayer coverage due to conformational changes of bacteria during the binding process, which may allow increased electron transfer [44]. Although the mechanism underlying the direct electron transfer from living organisms to the electrode is not well understood, it has been reported that the electrochemical response of bacteria is produced by extracellular electron transfer [48]. The redox mediator and Ab-SWCNT

conjugate modified electrode used also act as an electron shuttle enabling bacterial-electrode communication [49]. Ferricyanide is a hydrophilic redox mediator that has been reported to act as an electron acceptor for respiratory chain intermediates, such as cytochrome *a*, because it possesses stable Fe (II) and Fe (III) redox states. Cell redox sites or intermediates at the outer side of the cell membrane would lead to the reduction of ferricyanide [50]. Lastly, peak current enhancement may be due to sodium ions deposited on the bacterial cell wall [51]. When electric potential is applied during an electrochemical process, the adjoining environment becomes adverse to the cell, and the teichoic acid present in the bacterial cell wall [52] begins to accumulate cations present in the PBS buffer. These cations can enhance the conductivity of the whole cell, which can induce an increase in electrochemical peak current with increasing *S. aureus* concentration. Beyond these, as reported earlier, nanomaterials can also play an important role in increases or decreases of electrochemical responses with respect to increasing concentration of antigens [53–56].

To evaluate the performance of these immunosensors in food samples, 20% v/v milk in PBS was spiked with *S. aureus* at a concentration range of 10–10^7 CFU mL^{-1}. The immunosensor was able to detect *S. aureus* in these samples, as indicated by the change in peak current with increasing *S. aureus* concentration (Fig. 4b). This result demonstrates that these immunosensors can be used for the detection of *S. aureus* in actual food samples without the need of any pre-treatment.

The inset graphs in Fig. 4a and b show the linear relationship between the peak currents and log *S. aureus* concentration; the correlation coefficients (R^2) were 0.976 (a) in PBS and 0.987(b) in milk samples.

(a) In PBS, $I_p = 2.0639 \times \log(S.\ aureus\ \text{Concentration}) + 14.637$
(b) In Milk, $I_p = 1.7061 \times \log(S.\ aureus\ \text{Concentration}) + 24.579$

The limits of detection (LODs) were calculated from the calibration lines based on the signal to noise ratio method using a factor of 3. The standard deviations of the blank samples in PBS and milk were 0.51097 and 0.36160, respectively, and the estimated LODs were 15 CFU mL^{-1} in PBS and 13 CFU mL^{-1} in milk. Taking into account the reported infectious dose of *S. aureus* (10^3–10^8 CFU mL^{-1} [5]) together with these measured LODs and previous studies using other substrates, the current immunosensor appears to be highly sensitive. In fact, Braiek et al. [20] developed an immunosensor, using antibody-modified gold working electrodes in an electrochemical cell for the detection of *S. aureus* with an LOD of 10 CFU mL^{-1}. Chen et al. [57] also developed an electrochemical nanoparticle enzyme sensor for the detection of *S. aureus* in contaminated water, reporting an LOD of 100 CFU mL^{-1}.

3.4. Selectivity assay

Selectivity was tested by exposing these immunosensors to different bacteria (*E. coli* B, *B. subtilis*, and *S. epidermidis*) diluted in PBS (1×, pH 7.4) at a concentration range of 10^7 CFU mL^{-1} (Fig. 5a). The current change was much greater for *S. aureus* than for *E. coli* B, *B. subtilis*, and *S. epidermidis* at the same concentration, indicating good immunosensor specificity for *S. aureus*. Moreover, the peak currents of *E. coli* B, *B. subtilis*, and *S. epidermidis* were nearly the same as that observed following Ab-SWCNT conjugate functionalization on CPE.

3.5. Stability test

This paper-based electrochemical immunosensor was stored at 4 °C when not in use. The stability of the Ab-SWCNT bio-conjugate modified CPEs on the paper substrate was investigated using *S. aureus* (10^7 CFU mL^{-1}) for 5 weeks; no significant changes in current were observed (Fig. 5b). The bio-conjugates showed good stability: 1.94% (relative standard deviation [RSD]) on the same day, 3.73% RSD over 3 weeks, and 10.5% RSD after 5 weeks. Based on these stability data, we can conclude that these paper-based immunosensors were stable for storage and transportation.

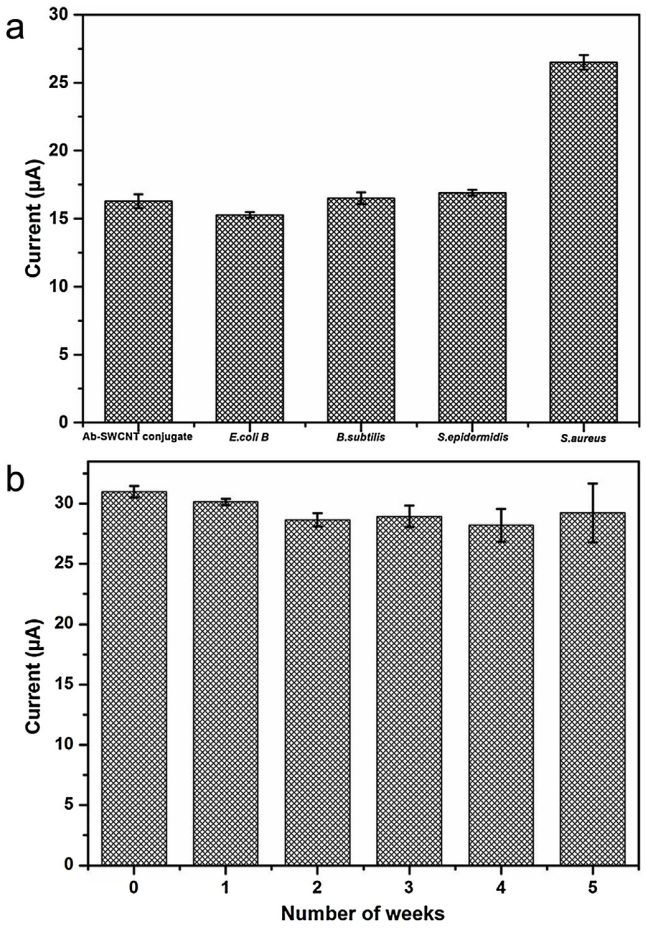


Fig. 5. Histogram depicting the (a) selectivity of the immunosensor for *S. aureus* vs. *E. coli* B, *B. subtilis*, and *S. epidermidis* and the (b) stability of immunosensor as a function of storage time over 5 weeks at 4 °C. The error bars represent the standard deviations of three measurements under optimal conditions.

4. Conclusion

In this study, a low-cost electrochemical immunosensor on paper substrate for label-free detection of *S. aureus* was developed using antibodies covalently conjugated to SWCNTs. The use of Ab-SWCNT conjugates can increase the robustness of a paper-based immunosensor by decreasing the functionalization steps because the wax-deposited hydrophobic barriers can be damaged by physical/chemical treatment such as bending, folding, functionalization, and immunoassay [29]. Covalent attachment of antibodies to SWCNTs reduced the functionalization steps on the paper surface to 1 h and the entire fabrication and immobilization process required approximately 2 h. This conjugation can also increase the stability of a biosensor by increasing the survival rate of biomolecules [30]. This paper-based electrochemical device was stable for over a month. A sensitive detection limit of 15 CFU mL^{-1} with a concentration range from 10 to 10^7 CFU mL^{-1} was obtained with good selectivity and a rapid detection time (~30 min). The work presented here can have potential to quantitatively monitor foodborne

bacteria with high sensitivity and selectivity in the field of food and environmental safety.

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (#2015R1A2A2A01006446).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.snb.2017.06.108.

References

[1] P. Leonard, S. Hearty, J. Brennan, L. Dunne, J. Quinn, T. Chakraborty, R. O'Kennedy, Advances in biosensors for detection of pathogens in food and water, Enzyme Microb. Technol. 32 (2003) 3–13, http://dx.doi.org/10.1016/S0141-0229(02)00232-6.
[2] B.J. McCabe-Sellers, S.E. Beattie, Food safety: emerging trends in foodborne illness surveillance and prevention, J. Am. Diet. Assoc. 104 (2004) 1708–1717, http://dx.doi.org/10.1016/j.jada.2004.08.028.
[3] J.A. Hennekinne, M.L. De Buyser, S. Dragacci, Staphylococcus aureus and its food poisoning toxins: characterization and outbreak investigation, FEMS Microbiol. Rev. 36 (2012) 815–836, http://dx.doi.org/10.1111/j.1574-6976.2011.00311.x.
[4] B. Kouidhi, T. Zmantar, H. Hentati, A. Bakhrouf, Cell surface hydrophobicity, biofilm formation, adhesives properties and molecular detection of adhesins genes in Staphylococcus aureus associated to dental caries, Microb. Pathog. 49 (2010) 14–22, http://dx.doi.org/10.1016/j.micpath.2010.03.007.
[5] H.C. Leggett, C.K. Cornwallis, S.A. West, Mechanisms of pathogenesis, infective dose and virulence in human parasites, PLoS Pathog. 8 (2012) e1002512, http://dx.doi.org/10.1371/journal.ppat.1002512.
[6] A. Chrouda, M. Braiek, K.B. Rokbani, A. Bakhrouf, A. Maaref, N. Jaffrezic-Renault, An immunosensor for pathogenic *Staphylococcus aureus* based on antibody modified aminophenyl-Au electrode, ISRN Electrochem. 2013 (2013) 1–9, http://dx.doi.org/10.1155/2013/367872.
[7] J. Schelin, N. Wallin-Carlquist, M.T. Cohn, R. Lindqvist, G.C. Barker, P. Rådström, The formation of Staphylococcus aureus enterotoxin in food environments and advances in risk assessment, Virulence 2 (2011) 580–592, http://dx.doi.org/10.4161/viru.2.6.18122.
[8] Scientific Criteria to Ensure Safe Food, National Academies Press, Washington, D.C, 2003, http://dx.doi.org/10.17226/10690.
[9] Y. Wang, J.K. Salazar, Culture-independent rapid detection methods for bacterial pathogens and toxins in food matrices, Compr. Rev. Food Sci. Food Saf. 15 (2015), http://dx.doi.org/10.1111/1541-4337.12175.
[10] J.W.F. Law, N.S. Ab Mutalib, K.G. Chan, L.H. Lee, Rapid methods for the detection of foodborne bacterial pathogens: principles, applications, advantages and limitations, Front. Microbiol. 5 (770) (2014), http://dx.doi.org/10.3389/fmicb.2014.00770.
[11] P.K. Mandal, A.K. Biswas, K. Choi, U.K. Pal, Methods for rapid detection of foodborne pathogens: an overview, Am. J. Food Technol. 6 (2011) 87–102, http://dx.doi.org/10.3923/ajft.2011.87.102.
[12] D.D. Liana, B. Raguse, J.J. Gooding, E. Chow, Recent advances in paper-based sensors, Sensors (Basel) 12 (2012) 11505–11526, http://dx.doi.org/10.3390/s120911505.
[13] K. Scida, B. Li, A.D. Ellington, R.M. Crooks, DNA detection using origami paper analytical devices, Anal. Chem. 85 (2013) 9713–9720, http://dx.doi.org/10.1021/ac402118a.
[14] P. von Lode, Point-of-care immunotesting: approaching the analytical performance of central laboratory methods, Clin. Biochem. 38 (2005) 591–606, http://dx.doi.org/10.1016/j.clinbiochem.2005.03.008.
[15] B. Liu, D. Du, X. Hua, X.-Y. Yu, Y. Lin, Paper-based electrochemical biosensors: from test strips to paper-based microfluidics, Electroanalysis 26 (2014) 1214–1223, http://dx.doi.org/10.1002/elan.201400036.
[16] C. Silveira, T. Monteiro, M. Almeida, Biosensing with paper-based miniaturized printed electrodes-a modern trend, Biosensors 6 (2016) 51, http://dx.doi.org/10.3390/bios6040051.
[17] P. Arora, A. Sindhu, N. Dilbaghi, A. Chaudhury, Biosensors as innovative tools for the detection of food borne pathogens, Biosens. Bioelectron. 28 (2011) 1–12, http://dx.doi.org/10.1016/j.bios.2011.06.002.
[18] S. Balasubramanian, I.B. Sorokulova, V.J. Vodyanoy, A.L. Simonian, Lytic phage as a specific and selective probe for detection of Staphylococcus aureus-a surface plasmon resonance spectroscopic study, Biosens. Bioelectron. 22 (2007) 948–955, http://dx.doi.org/10.1016/j.bios.2006.04.003.
[19] F. Tan, P.H.M. Leung, Z. Liu, Y. Zhang, L. Xiao, W. Ye, X. Zhang, L. Yi, M. Yang, A PDMS microfluidic impedance immunosensor for E. coli O157:H7 and Staphylococcus aureus detection via antibody-immobilized nanoporous membrane, Sensors Actuators B Chem. 159 (2011) 328–335, http://dx.doi.org/10.1016/j.snb.2011.06.074.
[20] M. Braiek, K.B. Rokbani, A. Chrouda, B. Mrabet, A. Bakhrouf, A. Maaref, N. Jaffrezic-Renault, An electrochemical immunosensor for detection of staphylococcus aureus bacteria based on immobilization of antibodies on self-assembled monolayers-functionalized gold electrode, Biosensors 2 (2012) 417–426, http://dx.doi.org/10.3390/bios2040417.
[21] V. Escamilla-Gómez, S. Campuzano, M. Pedrero, J.M. Pingarrón, Immunosensor for the determination of Staphylococcus aureus using a tyrosinase-mercaptopropionic acid modified electrode as an amperometric transducer, Anal. Bioanal. Chem. 391 (2008) 837–845, http://dx.doi.org/10.1007/s00216-007-1810-1.
[22] A. Abbaspour, F. Norouz-Sarvestani, A. Noori, N. Soltani, Aptamer-conjugated silver nanoparticles for electrochemical dual-aptamer-based sandwich detection of staphylococcus aureus, Biosens. Bioelectron. 68 (2015) 149–155, http://dx.doi.org/10.1016/j.bios.2014.12.040.
[23] A.W. Martinez, S.T. Phillips, M.J. Butte, G.M. Whitesides, Patterned paper as a platform for inexpensive, low-Volume, portable bioassays, Angew. Chemie. 119 (2007) 1340–1342, http://dx.doi.org/10.1002/ange.200603817.
[24] X. Li, J. Tian, G. Garnier, W. Shen, Fabrication of paper-based microfluidic sensors by printing, Colloids Surf. B. Biointerfaces 76 (2010) 564–570, http://dx.doi.org/10.1016/j.colsurfb.2009.12.023.
[25] G. Chitnis, Z. Ding, C.-L. Chang, C. a Savran, B. Ziaie, Laser-treated hydrophobic paper: an inexpensive microfluidic platform, Lab. Chip 11 (2011) 1161–1165, http://dx.doi.org/10.1039/c0lc00512f.
[26] W. Dungchai, O. Chailapakul, C.S. Henry, Electrochemical detection for paper-based microfluidics, Anal. Chem. 81 (2009) 5821–5826.
[27] Y. Lu, W. Shi, J. Qin, B. Lin, Fabrication and Characterization of Paper-Based Microfluidics Prepared in Nitrocellulose Membrane By Wax Printing, Anal. Chem. 82 (2010) 329–335.
[28] E. Carrilho, A.W. Martinez, G.M. Whitesides, Understanding wax printing: a simple micropatterning process for paper-based microfluidics, Anal. Chem. 81 (2009) 7091–7095, http://dx.doi.org/10.1021/ac901071p.
[29] Q. He, C. Ma, X. Hu, H. Chen, Method for fabrication of paper-based microfluidic devices by alkylsilane self-assembling and UV/O3-patterning, Anal. Chem. 85 (2013) 1327–1331, http://dx.doi.org/10.1021/ac303138x.
[30] K. Jiang, L.S. Schadler, R.W. Siegel, X. Zhang, H. Zhang, M. Terrones, Protein immobilization on carbon nanotubes via a two-step process of diimide-activated amidation, J. Mater. Chem. 14 (2004) 37, http://dx.doi.org/10.1039/b310359e.
[31] N. Shao, S. Lu, E. Wickstrom, B. Panchapakesan, Integrated molecular targeting of IGF1R and HER2 surface receptors and destruction of breast cancer cells using single wall carbon nanotubes, Nanotechnology 18 (2007) 315101, http://dx.doi.org/10.1088/0957-4484/18/31/315101.
[32] M.R. McDevitt, D. Chattopadhyay, B.J. Kappel, J.S. Jaggi, S.R. Schiffman, C. Antczak, J.T. Njardarson, R. Brentjens, D.A. Scheinberg, Tumor targeting with antibody-functionalized, radiolabeled carbon nanotubes, J. Nucl. Med. 48 (2007) 1180–1189, http://dx.doi.org/10.2967/jnumed.106.039131.
[33] P. Chakravarty, R. Marches, N.S. Zimmerman, A.D.-E. Swafford, P. Bajaj, I.H. Musselman, P. Pantano, R.K. Draper, E.S. Vitetta, Thermal ablation of tumor cells with antibody-functionalized single-walled carbon nanotubes, PNAS 25 (2008) 8697–8702.
[34] S. Pillai, S. Jain, S.R. Singh, D.W. Horn, V.A. Davis, M. Kumar Ram, Development of an antibody functionalized carbon nanotube biosensor for foodborne bacterial pathogens, J. Biosens Bioelectron. 11 (2012), http://dx.doi.org/10.4172/2155-6210.s11-002.
[35] C. Apetrei, I.M. Apetrei, J.A. de Saja, M.L. Rodriguez-Mendez, Carbon paste electrodes made from different carbonaceous materials: spplication in the study of antioxidants, Sensors 11 (2011) 1328–1344, http://dx.doi.org/10.3390/s110201328.
[36] A. Sharma, C.-H. Han, J. Jang, Rapid electrical immunoassay of the cardiac biomarker troponin I through dielectrophoretic concentration using imbedded electrodes, Biosens. Bioelectron. 82 (2016) 78–84, http://dx.doi.org/10.1016/j.bios.2016.03.056.
[37] A. Sharma, S. Hong, R. Singh, J. Jang, Single-walled carbon nanotube based transparent immunosensor for detection of a prostate cancer biomarker osteopontin, Anal. Chim. Acta 869 (2015) 68–73, http://dx.doi.org/10.1016/j.aca.2015.02.010.
[38] R. Singh, A. Sharma, S. Hong, J. Jang, Electrical immunosensor based on dielectrophoretically-deposited carbon nanotubes for detection of influenza virus H1N1, Analyst 139 (2014) 5415–5421, http://dx.doi.org/10.1039/c4an01335b.
[39] E. Venturelli, C. Fabbro, O. Chaloin, C. Ménard-Moyon, C.R. Smulski, T. Da Ros, K. Kostarelos, M. Prato, A. Bianco, Antibody covalent immobilization on carbon nanotubes and assessment of antigen binding, Small 7 (2011) 2179–2187, http://dx.doi.org/10.1002/smll.201100137.
[40] D.L. Pavia, G.M. Lampman, G.S. Kriz, J. Vyvyan, Introduction to Spectroscopy, 4th ed., Brooks/Cole, Cengage Learning Belmont, CA, 2009.
[41] H. Cells, L.G. Harris, S.J. Foster, R.G. Richards, L. Harris, An introduction to staphylococcus aureus, and techniques for identifying and quantifying s. aureus adhesins in relation to adhesion to biomaterials, Eur. Cell Mater. 4 (2002) 39–60.
[42] Z. Xu, N. Gao, H. Chen, S. Dong, Biopolymer and carbon nanotubes interface prepared by self-assembly for studying the electrochemistry of microperoxidase-11, Langmuir 21 (2005) 10808–10813, http://dx.doi.org/10.1021/la051445+.

[43] K.K.E.T. Mundelanji Vestergaard, An overview of label-free electrochemical protein sensors, Sensors (Basel) 7 (2007) 3442.
[44] J. Okuno, K. Maehashi, K. Kerman, Y. Takamura, K. Matsumoto, E. Tamiya, Label-free immunosensor for prostate-specific antigen based on single-walled carbon nanotube array-modified microelectrodes, Biosens. Bioelectron. 22 (2007) 2377–2381, http://dx.doi.org/10.1016/j.bios.2006.09.038.
[45] W. Hertl, Amperometmc immunoassays, Bioelectrochem. Bioenerg. 17 (1987) 89–100, http://dx.doi.org/10.1016/0302-4598(87)80010-7.
[46] S. Kumar, S. Kumar, C.M. Pandey, B.D. Malhotra, Conducting paper based sensor for cancer biomarker detection, J. Phys. Conf. Ser. 704 (2016) 012010, http://dx.doi.org/10.1088/1742-6596/704/1/012010.
[47] R. Singh, S. Hong, J. Jang, Label-free detection of influenza viruses using a reduced graphene oxide-based electrochemical immunosensor integrated with a microfluidic platform, Sci. Rep. 7 (2017) 42771, http://dx.doi.org/10.1038/srep42771.
[48] S. Han, X. Li, G. Guo, Y. Sun, Z. Yuan, Voltammetric measurement of microorganism populations, Anal. Chim. Acta 405 (2000) 115–121, http://dx.doi.org/10.1016/S0003-2670(99)00753-9.
[49] R.Y.A. Hassan, H.N.A. Hassan, M.S. Abdel-Aziz, E. Khaled, Nanomaterials-based microbial sensor for direct electrochemical detection of Streptomyces Spp, Sensors Actuators B Chem. 203 (2014) 848–853, http://dx.doi.org/10.1016/j.snb.2014.07.059.
[50] A. Bisschop, J. Boonstra, H.J. Sips, W.N. Konings, Respiratory chain linked ferricyanide reduction drives active transport in membrane vesicles from ba cill us sub tilis, FEBS Lett. 60 (1975).
[51] J.G. Swoboda, J. Campbell, T.C. Meredith, S. Walker, Wall teichoic acid function, biosynthesis and inhibition, Chembiochem 11 (2010) 35–45, http://dx.doi.org/10.1002/cbic.200900557.
[52] W.W. Navarre, O. Schneewind, Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope, Microbiol. Mol. Biol. Rev. 63 (1999) 174–229.
[53] S. Kumar, S. Kumar, S. Tiwari, S. Srivastava, M. Srivastava, B.K. Yadav, S. Kumar, T.T. Tran, A.K. Dewan, A. Mulchandani, J.G. Sharma, S. Maji, B.D. Malhotra, Biofunctionalized nanostructured zirconia for biomedical application: a smart approach for oral cancer detection, Adv. Sci. 2 (2015) 1500048, http://dx.doi.org/10.1002/advs.201500048.
[54] S. Kumar, J.G. Sharma, S. Maji, B.D. Malhotra, A biocompatible serine functionalized nanostructured zirconia based biosensing platform for non-invasive oral cancer detection, RSC Adv. 6 (2016) 77037–77046, http://dx.doi.org/10.1039/C6RA07392A.
[55] S. Kumar, J.G. Sharma, S. Maji, B.D. Malhotra, Nanostructured zirconia decorated reduced graphene oxide based efficient biosensing platform for non-invasive oral cancer detection, Biosens. Bioelectron. 78 (2016) 497–504, http://dx.doi.org/10.1016/j.bios.2015.11.084.
[56] S. Kumar, S. Kumar, S. Tiwari, S. Augustine, S. Srivastava, B.K. Yadav, B.D. Malhotra, Highly sensitive protein functionalized nanostructured hafnium oxide based biosensing platform for non-invasive oral cancer detection, Sensors Actuators B Chem. 235 (2016) 1–10, http://dx.doi.org/10.1016/j.snb.2016.05.047.
[57] J. Chen, Z. Jiang, J.D. Ackerman, M. Yazdani, S. Hou, S.R. Nugen, V.M. Rotello, Electrochemical nanoparticle-enzyme sensors for screening bacterial contamination in drinking water, Analyst 140 (2015) 4991–4996, http://dx.doi.org/10.1039/c5an00637f.

Biographies

Jyoti Bhardwaj received her B.Sc. in Biotechnology from Maharshi Dayanand University (MDU), Rohtak and M.Sc. in Biotechnology from Lovely Professional University (LPU), Jalandhar, India in 2010 and 2012, respectively. Then she worked as a research assistant in the field of Microbiology & Immunology from 2013 to 2014 at Chonnam National University (CNU), Yeosu, South Korea. She is now pursuing for a PhD degree in Department of Biomedical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, South Korea. Her main research interests are design, fabrication and development of paper based electrochemical sensors for detection of pathogenic micro-organisms.

Sivaranjani Devarakonda received her B.E degree from Osmania University, Hyderabad, India and M.S degree from Ulsan National Institute of Science and Technology in Mechanical Engineering in the years 2013 and 2016 respectively. Presently, she is doing her PhD in Mechanical Engineering of Iowa State University, Ames, Iowa starting from 2016. Her current research interest is in the study of flow through nanopores and microfabrication for possible applications in biosensors and drug delivery.

Suveen Kumar received his M.Sc. degree in Biotechnology and M.Tech. degree in Nanoscience and Technology from Guru Jambheshwar University of Science & Technology (GJUS&T), Hisar, Haryana, India in 2009 and 2011 respectively. Then he received his Ph.D. degree in Biotechnology from Delhi Technological University (DTU), Delhi, India in 2017. Presently, he is working as a Postdoctoral Fellow at the School of Mechanical, Aerospace and Nuclear Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Korea. He has published more than 12 research papers in peer reviewed international journals and his recent research interest is in electrochemical detection of airborne viruses.

Jaesung Jang received his B.S. and M.S. degrees in mechanical engineering from Pohang University of Science and Technology (POSTECH) in 1997 and 1999, respectively. Then he received his Ph.D. degree in mechanical engineering from Purdue University, West Lafayette, IN, in 2004. From 2004–2007, he worked as a postdoctoral research associate for the development of cantilever biosensors for detection of airborne viruses at the school of electrical and computer engineering, Purdue University, West Lafayette, IN. After serving as an assistant professor at Chung-Ang University, Seoul, Korea, for 2.5 years, he joined the faculty of the department of mechanical engineering at Ulsan National Institute of Science and Technology (UNIST), Ulsan, Korea, in 2010. He is now an associate professor and his research interests include collection, manipulation, concentration, separation, and detection of biological particles, especially for airborne biological particles.

Aquaculture 479 (2017) 134–141

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aquaculture

Effect of temperature on digestive physiology, immune-modulatory parameters, and expression level of Hsp and LDH genes in *Catla catla* (Hamilton, 1822)

JaiGopal Sharma[a], Samar Pal Singh[b], Rina Chakrabarti[b,*]

[a] Department of Biotechnology, Delhi Technological University, Bawana Road, Delhi 110042, India
[b] Aqua Research Lab, Department of Zoology, University of Delhi, Delhi 110 007, India

ARTICLE INFO

Keywords:
Catla
Heat shock proteins
Lysozyme
Myeloperoxidase
Nitric oxide synthase
Temperature stress

ABSTRACT

Fish face the challenge of diurnal and seasonal fluctuations of temperature in culture and it is essential to understand their physiological responses to thermal stress. Indian major carp *Catla catla*, catla (7.52 ± 0.22 g) were acclimated at 28 ± 1 °C for 21 days and then exposed at six different temperatures: 10, 15, 20, 25, 28 and 35 °C. Blood and tissue samples were collected after 12 h and 7 days of reaching the assigned temperature. Significantly ($p < 0.05$) higher mortality of fish was recorded at 10 °C. Exposure of catla below 25 °C affected food consumption and digestive enzyme activities. It caused immunosupression as significantly ($p < 0.05$) lower serum lysozyme, myeloperoxidase and nitric oxide synthase levels were found in fish. Significantly ($p < 0.05$) higher thiobarbituric acid reactive substances indicated higher lipid peroxidation in catla exposed at 10–15 °C compared to other treatments. The expression of heat shock protein (Hsp) genes Hsp60, Hsp70 and Hsp90 and lactate dehydrogenase, LDH-A was evaluated in hepatopancreas and muscles. Hsp60 was sensitive towards lowering of temperature, whereas Hsp70 was responsive at highest temperature in both tissues. This study highlighted the impacts of cold environmental conditions on the physiology of this species.

1. Introduction

Successful aquaculture depends on the environment in which the organism is cultured. An optimum environment must be maintained over the culture period to achieve maximum production. Among various abiotic factors temperature has been considered as a master abiotic factor. Water temperature influences other environmental factors viz. dissolved oxygen, photosynthesis and aerobic respiration (Zhou et al., 2010). Water temperature is an important factor for poikilothermic fish. It influences the rate of feeding, metabolism and growth of fish. Extreme temperature can alter the enzymatic activities. Rapid change in water temperature may cause thermal shock. Temperature beyond the optimal limit of a particular species may cause a variety of pathophysiological disturbances that can lead to the death of the species (Dalvi et al., 2009). Temperature can influence the innate immune system and the immune competence of fish (Pérez-Casanova et al., 2008; Singh et al., 2013; Dittmar et al., 2014).

Catla catla, catla is one of the economically important carp species used in composite culture throughout India. Catla is a surface feeder and zooplanktivore species (FAO, 2017a, http://www.fao.org/fishery/affris/species-profiles/catla/natural-food-and-feeding-habits/en/). Catla, along with other carps, contributes approximately 5.4 million - tonnes to the total aquaculture production. India also dominates the global catla production by contributing 71% to the global total of 3.87 million tonnes (FAO, 2012). Wide temperature fluctuation is recorded in different parts of India in various seasons. Therefore, this is essential to understand the impact of physiological responses of cultivable species to various temperatures that fish may come across at different hours of the day and in various seasons. Exposure of organisms to heat shock shows harmful effects on the internal cell organization beyond individual proteins unfolding (Richter et al., 2010; Toivola et al., 2010). Heat shock proteins (Hsps) serve as potential biomarkers to understand the response of fish towards temperature stress. Hsp plays significant role in folding, fidelity and translocation of proteins; it inhibits faulty interactions and thus, gives protection to vital cellular functions. Lactate dehydrogenase (LDH), the glycolytic enzyme has pre-eminent importance in energy metabolism. The activity and the functional properties of this enzyme reflect the amplitude for anaerobic energy production, and by that, the level of resistance to oxygen deficiency during hypoxia and thermal stress (Somero, 1998; Pörtner, 2002). The present investigation aimed to understand the impact of various temperatures on the digestive physiology, immune system, lipid

* Corresponding author.
E-mail address: aquaresearchlab@yahoo.co.in (R. Chakrabarti).

http://dx.doi.org/10.1016/j.aquaculture.2017.05.031
Received 3 February 2017; Received in revised form 17 May 2017; Accepted 24 May 2017
Available online 25 May 2017
0044-8486/ © 2017 Published by Elsevier B.V.

peroxidation and expression of selected genes of *Catla catla*.

2. Materials and methods

2.1. Exposure of fish to various temperatures

Indian major carp *Catla catla*, catla (7.52 ± 0.22 g) were randomly distributed in glass aquaria (50 l each) and acclimated for 21 days at ambient temperature of 28 ± 1 °C. The stocking density was 20 fish aquarium^{-1}. Each aquarium was connected with a filtration unit (Sera fil bioactive 130, Germany) for recirculation and water quality maintenance; aquaria were covered with transparent fiber glass sheet to avoid escape of fish, this also helped to maintain the desirable temperature in the aquaria. The use of a filter and aerator assisted the maintenance of total fish biomass (3 g l^{-1}) in the culture units. Fish were exposed to one of six different temperatures: 10, 15, 20, 25, 28 and 35 °C. The acclimation temperature was considered as control. Three replicates were used for each temperature. The cooling (Hailea Chiller HC-300A, China) or heating (Hicon, India) unit was connected with each aquarium for the maintenance of desirable temperature. Each experimental temperature was achieved at a rate of change of 2 °C day^{-1} starting from acclimation temperature. This required 1.5–9 days to reach the assigned temperature. Samples were collected after 12 h and 7 days of reaching the assigned temperature to evaluate the effect of immediate stress and chronic stress, respectively on the physiology of catla.

Dissolved oxygen level was maintained 5 mg l^{-1} with the help of aerator; photoperiod was 12 h: 12 h. Pelleted feed (40% protein) was prepared in the laboratory using dry fish powder, wheat flour, cod liver oil and vitamin mineral premixes. Fish were fed at the rate of 5% of body weight per day throughout the study period. Water quality parameters like temperature, pH and dissolved oxygen were monitored regularly with a portable meter (HQ40D Multimeter, Hach, USA) using probes. Survival rate of fish was recorded.

2.2. Food consumption rate

Food consumption rate was calculated for individual fish after each meal. Fish were allowed to feed for 1 h. After 1 h of feeding, unconsumed food was siphoned from individual aquarium. The uneaten food can be easily differentiated from faeces by their physical textures. The collected food was dried, and the weight of the dry food was recorded. The amount of food consumed during each meal was calculated as the difference between the mass of the food offered and that of the remained food in the aquarium after drying. It was expressed as g fish^{-1}.

2.3. Blood and tissue collection

Fish were euthanized with tricaine methanesulphonate (MS-222 Sigma-Aldrich, USA). Fish were immersed in solution of MS-222 (250 mg l^{-1}) for 10 min following cessation of opercular movement. Blood was collected from caudal vein of individual fish using disposable syringe. Pooled blood samples were used. Blood samples were collected in two different ways. For the study of respiratory brust activity, blood sample was taken in a micro-centrifuge tube coated with anticoagulant ethylenediamine tetraacetic acid (EDTA). Simultaneously, blood sample was collected in a micro-centrifuge tube without EDTA, allowed to clot for 1 h at room temperature and stored at 4 °C overnight. Then centrifuged at 1500 × *g* for 10 min. The serum was kept at 20 °C for further study.

Tissues, muscles and digestive tract were collected aseptically, frozen immediately with liquid nitrogen and stored at − 80 °C until used for analysis. For gene expression study, hepatopancreas (the diffused liver and pancreas, the active site for metabolic activities, hepatic tissues are playing major role in detoxification) and muscle (the major component of fish body) were selected. Tissues were collected aseptically and stored immediately in RNA*later* solution (Sigma-Aldrich) at − 20 °C before extraction of total RNA.

2.4. Biochemical assays

2.4.1. Digestive enzymes

The digestive tract was collected from individual fish, cleaned with chilled distilled water and blot dried. Tissue was homogenised with pre-chilled distilled water (1:10, w/v) and centrifuged at 10,000 × *g* at 4 °C for 10 min. The supernatant was used for various enzyme assays. Total soluble protein was estimated (Bradford, 1976) using bovine serum albumin as standard.

Amylase activity was evaluated by measuring the increase in reducing power of buffered starch solution with 3, 5-dinitro salicylic acid (Bernfeld, 1955). The absorbance was recorded at 540 nm using UV-1601 Spectrophotometer (Shimadzu, Japan). Specific amylase activity was expressed as mg maltose liberated mg protein^{-1} h^{-1} at 37 °C. Total protease activity was measured following the method of García-Carreño (1992) using 1% azocasein as substrate. The activity was expressed as Units mg protein^{-1} min^{-1}. Trypsin and chymotrypsin activities were estimated following the method of Erlanger et al. (1961) using *N*-*α*-benzoyl-DL-arginine-*p*-nitroanilide (BAPNA) and succinyl-alanine-alanine-proline-phenylalanine-*p*-nitroanilide (SAPNA, Sigma) as substrate, respectively. Lipase activity was assayed following the principle of colorimetric estimation of *p*-nitrophenol released by enzymatic hydrolysis of *p*-nitrophenyl palmitate and the optical density was recorded at 410 nm (Winkler and Stuckmann, 1979).

2.4.2. Immunological assays

A turbidometric assay utilizing lyophilized *Micrococcus lysodeikticus* cells was used to determine serum lysozyme (Siwicki, 1989). Absorbance was recorded at 450 nm using microplate reader (BioTek Synergy HT, USA) and expressed as mg ml^{-1}. Lyophilized hen's egg white lysozyme (Genei, India) was used as standard. Respiratory burst activity was measured using fresh blood with EDTA (Anderson and Siwicki, 1995). The extent of nitroblue tetrazolium reduced was measured at 540 nm. Dimethyl formamide was used as the blank. The myeloperoxidase activity of serum was measured according to Quade and Roth (1997). Hank's balanced salt solution was used as blank. Optical density was measured at 450 nm. Nitric oxide synthase in muscles was estimated (Lee et al., 2003). The optical density was recorded at 540 nm using microplate reader and expressed as mmol of nitrite mg tissue^{-1}.

2.4.3. Lipid peroxidation

Thiobarbituric acid reactive substances (TBARS) content in muscles was measured following the method of Ohkawa et al. (1979), in which malondialdehyde (MDA), the end product of lipid peroxidation reacts with thiobarbituric acid (TBA). Optical density was recorded at 532 nm and expressed as μmol MDA mg protein^{-1}.

2.5. Gene expression study

2.5.1. RNA isolation and first strand cDNA synthesis

Total genomic RNA was extracted from muscles and hepatopancreas with TRIzol reagent (Invitrogen, USA) following manufacturer's protocol. The concentration of RNA was measured using Nanodrop ND-1000 (Thermo Scientific, USA). RNA integrity was checked by observing band intensity on agarose gel (1%). For the synthesis of first strand cDNA, 1 μg of RNA was treated with 1 U of DNase I (Sigma-Aldrich, USA). This treatment helps to avoid DNA contamination. DNase treated RNA was reverse transcribed using ProtoScript-M-MuLV 1st strand cDNA synthesis kit (New England Biolabs, USA) following manufacturer's instructions. β-actin, the constitutively expressed housekeeping gene was used for sample normalization and as positive control. The

Table 1
Target genes and sequences of primers used to investigate mRNA levels.

Target gene	Primer	Primer sequence (5′ → 3′)	Accession number
β-actin	β-actin Fw	TCCTGACCGAGAGAGGCTAC	JQ991014.1
	β-actin Rv	CTGCTCGAAGTCAAGAGCCA	
Hsp60	Hsp60 Fw	TGCTGTTGCAAAGGAGGGAT	KC599205.1
	Hsp60 Rv	TGACCGGCTTGGAGAGTTTC	
Hsp70	Hsp70 Fw	CATGTGAGCGAGCCAAGAGA	KC599207.1
	Hsp70 Rv	TTCAAAGCGAGCTCTGGTGA	
Hsp90	Hsp90 Fw	TCGAACTCATCCCCGATCTGA	KC800801.1
	Hsp90 Rv	CTTTGTGCCGGATTTAGCGAT	
LDH-A	LDH-A Fw	GAAACGAGCAGAAAGGAGAAGCTAT	KC887544.1
	LDH-A Rv	TTGGCGGTCACACTGTAGTCTT	

PCR-amplification of β-actin was carried out for the confirmation of cDNA synthesis.

2.5.2. Real-time PCR analysis

Quantitative real-time PCR was carried out in ViiA-7™ real-time PCR system (Applied Biosystems, USA) using MicroAmp optical 96-well reaction plates with adhesive optical cover. All the primers for different genes were designed using Primer 3 software (Table 1). The cDNA was diluted in 1:3 with nuclease free water. Amplifications were performed in 10 µl reaction volume. Each reaction consisted of 1 µl of diluted cDNA (10 ng), 0.25 µl each of forward (FW) and reverse (RV) primer (concentration 2.5 µmol each), 5 µl of 2× Power SYBR® green PCR master mix (Invitrogen, USA) and 3.5 µl of PCR-grade water. The PCR amplification was performed in duplicate for each sample under following conditions: initial denaturation at 95 °C for 10 min, followed by 40 cycles of 95 °C/15 s, 60 °C/min. In negative control (NTC), water was added instead of cDNA. The PCR efficiencies were estimated by analysis of serial dilutions of cDNA. The efficiencies close to 100% were used in the $2^{-\Delta\Delta CT}$ method for relative gene expression calculation (Livak and Schmittgen, 2001). The relative expressions of target genes were normalized to the expression of β-actin in the same sample. Specificity of the q-PCR amplifications was determined using melt curves with a temperature range of 60–95 °C. The real-time PCR products were loaded on the agarose gel (1%) to confirm the correct size and single band amplification.

2.6. Statistical analysis

Results were expressed as mean ± SE. ANOVA two factor analysis was used to see the interaction between the treatments and time (SPSS software 19.0). Statistical significance was accepted at $p < 0.05$ level.

3. Results

3.1. Mortality of fish

The mortality of fish was recorded at 10 °C. In this treatment, 52.0% fish died within 12 h of exposure. Therefore, no fish was available for second sampling, scheduled after 7 days of reaching 10 °C temperature. There was no mortality of fish in other treatments.

3.2. Food consumption rate

There was no significant ($p > 0.05$) difference in food consumption rate (0.19 g fish^{-1}) of catla maintained in various aquaria during acclimation period (28 °C). Reduced food consumption rate was recorded in fish exposed at 20, 15 and 10 °C (Figs. 1a, b, c) compared to the fish cultured at 28 °C. In all these treatments, food consumption rate started decreasing as temperature dropped to 22 °C from 28 °C. A 59, 58 and 26% reduced food consumption rate was recorded in catla exposed at 10, 15 and 20 °C, respectively compared to the fish cultured at acclimation temperature (28 °C). There was no significant ($p > 0.05$) change in food consumption rate of catla exposed at 25 and 35 °C compared to the control one.

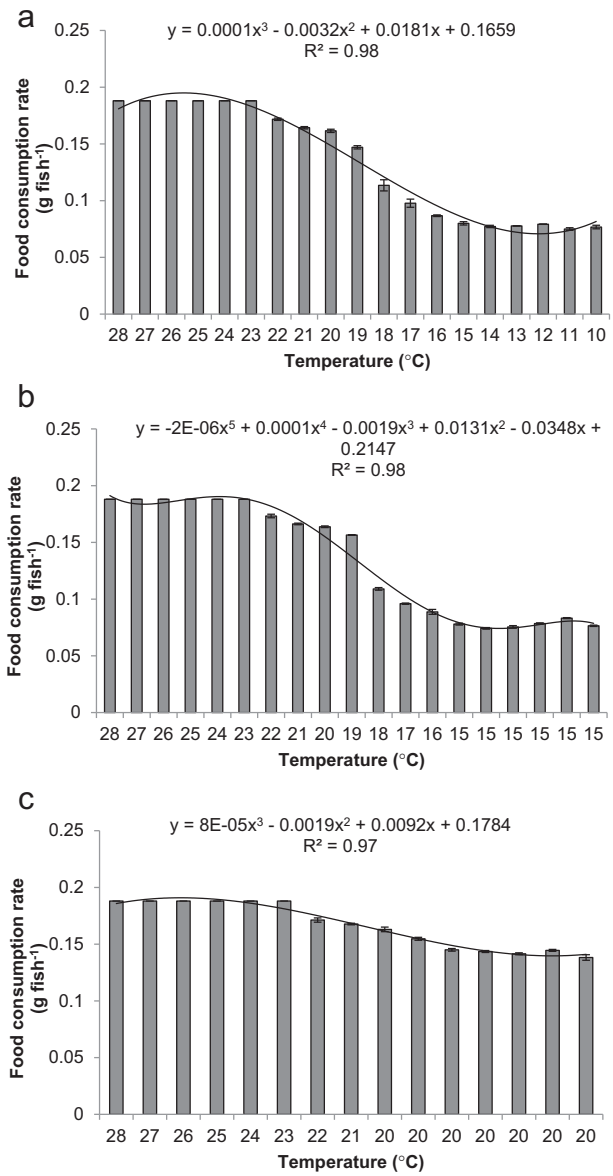


Fig. 1. Food consumption rate of *Catla catla* exposed at (a) 10, (b) 15 and (c) 20 °C.

3.3. Biochemical assays

3.3.1. Digestive enzymes

Amylase activity was significantly ($p < 0.05$) higher in catla maintained at 25 and 28 °C compared to other treatments after 12 h of exposure. After 7 days, amylase activity was significantly ($p < 0.05$) higher at 28 and 35 °C compared to other treatments. The activity was significantly ($p < 0.05$) lower after 7 days in catla exposed at 25 and 28 °C compared to the respective sample of 12 h. Total protease activity was significantly ($p < 0.05$) higher at 25 and 28 °C compared to others after 12 h of exposure. The activity was significantly ($p < 0.05$) higher at 28 °C compared to other treatments after 7 days of exposure. The activity reduced significantly ($p < 0.05$) in all treatments (except 35 °C) after 7 days compared to the 12 h exposed catla in the respective treatment. Trypsin activity was significantly ($p < 0.05$) higher at 25 and 28 °C compared to the other treatments after 12 h of exposure. After 7 days, activity was significantly ($p < 0.05$) higher in 28 and

Table 2
Digestive enzyme activities in *Catla catla* exposed at different temperatures. Results are expressed as mean ± SE ($n = 3$). Values with different superscripts in same column exposed for same duration are significantly ($p < 0.05$) different.

	Sampling	10 °C	15 °C	20 °C	25 °C	28 °C	35 °C
Amylase (mg maltose mg protein^{-1} h^{-1})	12 h	0.49 ± 0.02c	0.81 ± 0.05b	0.90 ± 0.03b	1.18 ± 0.15a	1.24 ± 0.11a	0.88 ± 0.02b
	7 days	–	0.58 ± 0.08c	0.75 ± 0.05b	0.78 ± 0.05b	0.89 ± 0.02a	0.87 ± 0.02a
Total protease activity (U mg protein^{-1} min^{-1})	12 h	0.19 ± 0.02d	0.41 ± 0.03c	1.13 ± 0.22b	2.57 ± 0.20a	2.66 ± 0.20a	1.32 ± 0.20b
	7 days	–	0.13 ± 0.02e	0.50 ± 0.04d	0.95 ± 0.08c	1.75 ± 0.10a	1.40 ± 0.08b
Specific trypsin activity (U mg protein^{-1} min^{-1})	12 h	0.09 ± 0.01e	0.12 ± 0.01d	0.20 ± 0.03c	0.33 ± 0.03a	0.38 ± 0.03a	0.28 ± 0.02b
	7 days	–	0.11 ± 0.02c	0.22 ± 0.02b	0.26 ± 0.03b	0.43 ± 0.02a	0.46 ± 0.04a
Specific chymotrypsin activity (U mg protein^{-1} min^{-1})	12 h	0.15 ± 0.01e	0.23 ± 0.02d	1.33 ± 0.14c	2.34 ± 0.12a	2.77 ± 0.32a	1.75 ± 0.23b
	7 days	–	0.19 ± 0.02e	0.98 ± 0.04d	1.28 ± 0.19c	2.26 ± 0.09a	1.92 ± 0.08b
Lipase (μM *p*-nitrophenol ml^{-1} min^{-1})	12 h	2.57 ± 0.24c	2.43 ± 0.23c	5.33 ± 0.45b	6.66 ± 0.53a	7.18 ± 0.30a	6.71 ± 0.72a
	7 days	–	2.06 ± 0.14c	5.75 ± 0.35b	6.15 ± 0.24b	6.05 ± 0.71b	9.62 ± 0.26a

35 °C compared to the other treatments. Chymotrypsin activity was significantly ($p < 0.05$) higher at 25 and 28 °C after 12 h and at 28 °C after 7 days of exposure compared to the other treatments. Lipase activity was significantly ($p < 0.05$) higher at 25–35 °C compared to 10–20 °C treatments after 12 h of exposure. After 7 days, activity was significantly ($p < 0.05$) higher at 35 °C compared to other treatments. After 7 days, the activity increased significantly ($p < 0.05$) at 35 °C compared to the fish exposed in the same treatment after 12 h (Table 2). Enzyme activity showed a direct relationship with the decrease of temperature in the culture system. This might be reflected in the less food consumption of catla. The two-way analysis of variance (ANOVA) of all digestive enzymes showed significant ($p \leq 0.05$) interaction between the temperature and time.

3.3.2. Immunological assays

Serum lysozyme level was significantly ($p < 0.05$) higher in catla exposed at 35 °C compared to other treatments after 12 h of exposure. Significantly ($p < 0.05$) lower level of lysozyme was found at 10 °C compared to other treatments. After 7 days, lysozyme level was significantly ($p < 0.05$) higher at 28 and 35 °C compared to other treatments. After 7 days, lysozyme level reduced significantly ($p < 0.05$) at 15, 25 and 35 °C compared to 12 h exposed catla in the respective treatment. There was no significant ($p > 0.05$) difference in respiratory burst activity among treatments after 12 h. After 7 days, the activity was significantly ($p < 0.05$) higher at 28 °C compared to others. Myeloperoxidase activity was significantly ($p < 0.05$) higher at 25 and 28 °C compared to the catla exposed to other treatments after 12 h and 7 days of exposure. Nitric oxide synthase (NOS) level was significantly ($p < 0.05$) higher at 28 °C compared to others in both samples. After 7 days, NOS level reduced significantly ($p < 0.05$) at 15, 25 and 28 °C compared to 12 h exposed catla (Table 3). The two-way analysis of variance (ANOVA) of lysozyme showed a significant ($p < 0.05$) interaction between the temperature and time; there was no significant ($p > 0.05$) interaction for other parameters like, myeloperoxidase, nitric oxide synthase and respiratory burst activity.

3.3.3. Lipid peroxidation

Thiobarbituric acid reactive substances (TBARS) level was significantly ($p < 0.05$) higher at 10 and 15 °C compared to others after 12 h and 7 days, respectively. TBARS level increased significantly ($p < 0.05$) after 7 days compared to 12 h exposed catla regardless of treatment (Table 3). The two-way analysis of variance (ANOVA) of TBARS showed significant ($p < 0.01$) interaction between the temperature and time.

3.4. Gene expression

3.4.1. Heat shock proteins Hsp60, Hsp70 and Hsp90

The expression of Hsp60, Hsp70 and Hsp90 was evaluated in muscles and hepatopancreas. In muscles, the expression of Hsp60 was significantly ($p < 0.05$) higher in all treatments compared to 28 °C after 12 h of exposure. After 7 days, it was significantly ($p < 0.05$) down-regulated at 25 and 35 °C and up-regulated at 15 °C compared to the control one (Fig. 2a). In hepatopancreas, the expression of Hsp60 was significantly ($p < 0.05$) up-regulated at 25–10 °C compared to 28 °C after 12 h. After 7 days, the expression was significantly ($p < 0.05$) up-regulated at 15 and 35 °C compared to 28 °C (Fig. 2b).

In muscles, the expression of Hsp70 was significantly ($p < 0.05$) higher at 10 °C compared to others after 12 h of exposure. After 7 days, the expression was induced in muscles at 15, 20 and 35 °C compared to 28 °C treatment (Fig. 3a). In hepatopancreas, the expression was significantly ($p < 0.05$) induced at 10 °C compared to catla at 28 °C after 12 h of exposure. The expression was significantly ($p < 0.05$) down-regulated at 25 and 20 °C compared to fish at 28 °C. After 7 days, the expression was significantly ($p < 0.05$) induced at 15, 20 and 35 °C compared to 28 °C treatment (Fig. 3b). Highest expression was found at 35 °C.

In muscles, Hsp90 expression was significantly ($p < 0.05$) up-regulated at 25 and 20 °C and down regulated at 15 and 10 °C, compared to catla at 28 °C after 12 h of exposure (Fig. 4a). The expression was significantly ($p < 0.05$) higher at 35 °C compared to the control one. In hepatopancreas, the expression of Hsp90 was

Table 3
Effect of various temperatures on the immune system and lipid peroxidation of *Catla catla* after 12 h and 7 days exposure. Results are expressed as means ± SE ($n = 3$). Values with different superscripts in same column exposed for same duration are significantly ($p < 0.05$) different.

	Sampling	10 °C	15 °C	20 °C	25 °C	28 °C	35 °C
Lysozyme (mg ml^{-1})	12 h	1.54 ± 0.17d	2.15 ± 0.05c	2.29 ± 0.10b	2.41 ± 0.03b	2.45 ± 0.05b	2.87 ± 0.10a
	7 days	–	1.83 ± 0.13c	2.20 ± 0.08b	2.14 ± 0.1b	2.45 ± 0.12a	2.47 ± 0.10a
Respiratory burst activity (O.D. at 540 nm)	12 h	0.31 ± 0.02a	0.31 ± 0.03a	0.34 ± 0.05a	0.34 ± 0.04a	0.40 ± 0.08a	0.35 ± 0.04a
	7 days	–	0.37 ± 0.01b	0.39 ± 0.02b	0.38 ± 0.04b	0.52 ± 0.05a	0.33 ± 0.01c
Myeloperoxidase activity (O.D. at 450 nm)	12 h	2.14 ± 0.03c	2.28 ± 0.04b	2.40 ± 0.05b	2.56 ± 0.10a	2.55 ± 0.05a	2.31 ± 0.05b
	7 days	–	2.35 ± 0.06b	2.40 ± 0.04b	2.55 ± 0.05a	2.65 ± 0.07a	2.30 ± 0.05b
Nitric oxide synthase (mmol mg tissue^{-1})	12 h	7.90 ± 0.70d	10.55 ± 0.16c	10.55 ± 0.37c	13.71 ± 0.59b	17.95 ± 0.81a	9.94 ± 0.48c
	7 days	–	7.96 ± 0.50c	9.07 ± 0.93b	9.83 ± 0.37b	14.2 ± 0.37a	8.79 ± 0.59b
Lipid peroxidation (μmol MDA mg protein^{-1})	12 h	20.9 ± 2.81a	15.9 ± 1.88b	13.6 ± 2.46b	13.8 ± 1.58b	10.9 ± 1.71c	14.9 ± 1.72b
	7 days	–	40.0 ± 4.56a	26.3 ± 3.64c	26.8 ± 2.94c	17.9 ± 1.18d	33.3 ± 2.0b

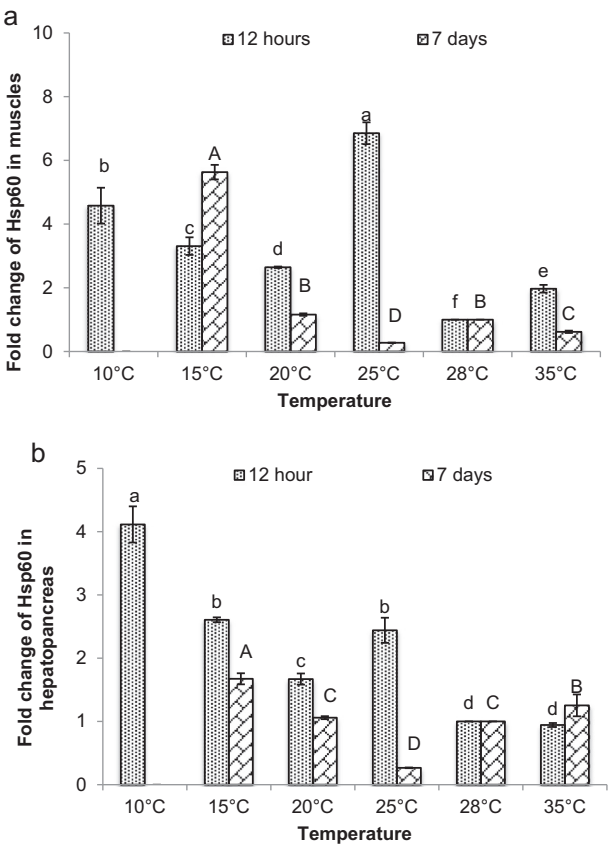


Fig. 2. Expression of Hsp60 in (a) muscles and (b) hepatopancreas of *Catla catla* exposed at various temperatures. Samples were collected after 12 h and 7 days of exposure. Bars with different superscripts are significantly ($p < 0.05$) different ($n = 6$). There is no data in 10 °C treatment after 7 days as all fish died after 12 h of exposure.

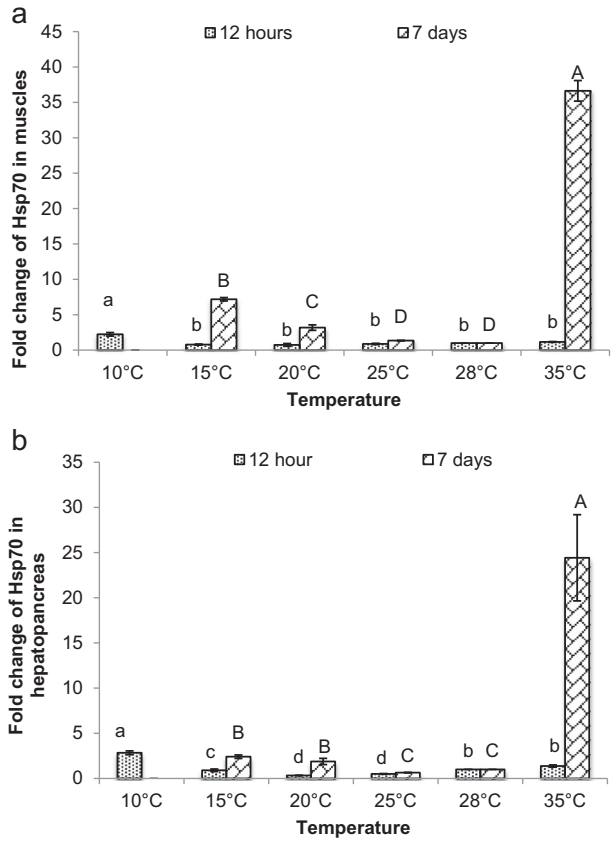


Fig. 3. Expression of Hsp70 in (a) muscles and (b) hepatopancreas of *Catla catla* exposed at various temperatures. Samples were collected after 12 h and 7 days of exposure. Bars with different superscripts are significantly ($p < 0.05$) different ($n = 6$). There is no data in 10 °C treatment after 7 days as all fish died after 12 h of exposure.

significantly ($p < 0.05$) up-regulated in 15 and 35 °C treatments compared to 28 °C treatment, after 12 h of exposure. Similar trend of Hsp90 expression in hepatopancreas was found after 7 days of exposure. Expression was significantly induced at 15 and 35 °C compared to 28 °C (Fig. 4b).

The two-way analysis of variance (ANOVA) of all heat shock proteins showed significant ($p < 0.01$) interaction between the temperature and time for hepatopancreas and muscles.

3.4.2. LDH-A

In muscles, the expression of LDH-A was significantly ($p < 0.05$) up-regulated in 10–25 °C treatments compared to the control one after 12 h of exposure. The expression was significantly ($p < 0.05$) up-regulated at 15, 20 and 35 °C and down-regulated at 25 °C compared to 28 °C treatment after 7 days of exposure (Fig. 5a). In hepatopancreas, the expression of LDH-A was significantly ($p < 0.05$) up-regulated in all treatments (except 35 °C) compared to catla exposed at 28 °C after 12 h. After 7 days, the expression of LDH-A was significantly ($p < 0.05$) up-regulated in 15, 20 and 35 °C compared to 28 °C treatment (Fig. 5b). The two-way analysis of variance (ANOVA) of LDH-A for hepatopancreas and muscles showed significant ($p < 0.01$) interaction between the temperature and time.

4. Discussion

Water temperature plays pivotal role in the lives of aquatic organisms. In the natural water bodies and in the aquaculture systems, fishes try to adjust themselves with the fluctuations of temperature. This may affect the overall physiology that leads to poor survival rate of the species. In the present study, mortality of catla occurred at lowest temperature 10 °C. The minimum temperature tolerance limit of catla is ~14 °C. The highest temperature 35 °C used in the present study was 3 °C higher compared to the optimum range of temperature 25–32 °C for catla (FAO, 2017b, http://www.fao.org/fishery/culturedspecies/Catla_catla/en). In cold and temperate regions, the temperature-mortality rises as temperatures fall below or rise above this "point of maximum comfort".

Food consumption rate of catla reduced as the water temperature decreased (below 25 °C). Reduction in temperature also affected the digestive enzyme activities. Poor digestibility resulted in reduced appetite in the cultured species. This was co-related with the reduced swimming activity of the fish. Low temperatures have a significant negative impact on fish appetite, growth and activity (Brown et al., 1989; Clark et al., 1995; Bendiksen et al., 2002). Atlantic salmon *Salmo salar* responded to the temperature fluctuation with reduced feed intake (Folkedal et al., 2010). Catla exposed at 10–20 °C showed reduced digestive enzymes activities. In common carp *Cyprinus carpio*, feed conversion efficiency decreased with low temperatures (Goolish and Adelman, 1984), which might be result of enzyme kinetics (Singh et al., 2009). In catfish *Clarias batrachus* food consumption and digestive enzyme activities were affected by lowering of temperature (Ahmad et al., 2014). Temperature beyond optimum limits adversely affected the health of the animals due to metabolic stress, which suppressed growth, reproduction and immune capacity (Cnaani, 2006). Reduced food consumption and digestive enzyme activity are sensitive indicators of stress in fish.

The study of biochemical parameters in fishes revealed information on the physiological changes occurring in response to decrease or increase of water temperature from the acclimation temperature. In catla, lysozyme level reduced when the water temperature decreased

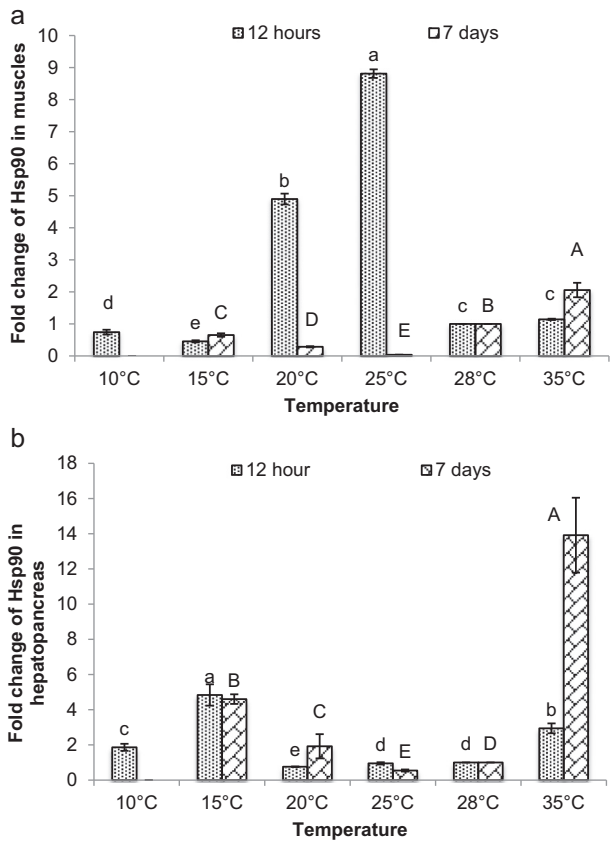


Fig. 4. Expression of Hsp90 in (a) muscles and (b) hepatopancreas of *Catla catla* exposed at various temperatures. Samples were collected after 12 h and 7 days of exposure. Bars with different superscripts are significantly ($p < 0.05$) different ($n = 6$). There is no data in 10 °C treatment after 7 days as all fish died after 12 h of exposure.

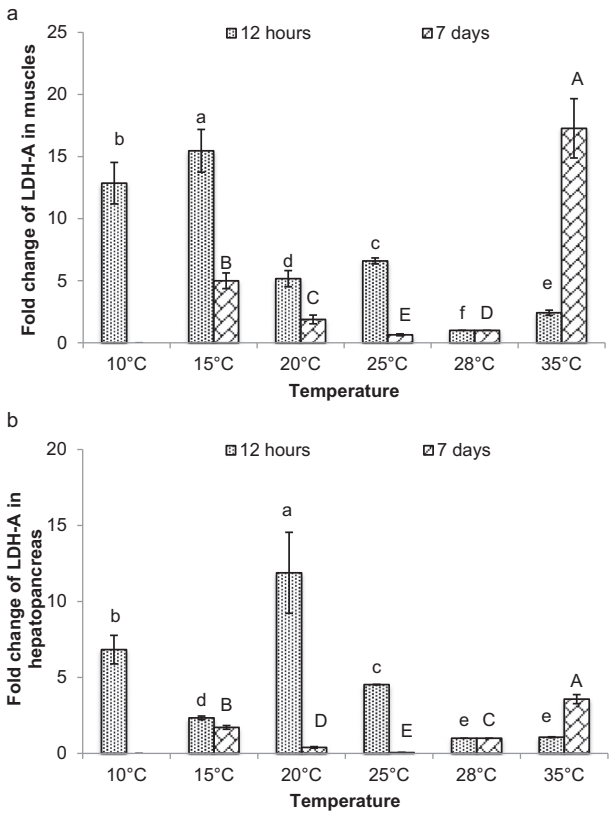


Fig. 5. Expression of LDH-A in (a) muscles and (b) hepatopancreas of *Catla catla* exposed at various temperatures. Samples were collected after 12 h and 7 days of exposure. Bars with different superscripts are significantly ($p < 0.05$) different ($n = 6$). There is no data in 10 °C treatment after 7 days as all fish died after 12 h of exposure.

from the acclimation one. A direct relationship was found between the water temperature and lysozyme level. Fish lysozymes possess a high potential for bacteriocidal or bacteriolytic activity and play an important role in the bio-defence system against Gram-positive and Gram-negative bacteria (Saurabh and Sahoo, 2008). In the present study, respiratory burst activity was not responsive towards immediate temperature stress; the activity decreased significantly with the lowering of temperature after 7 days of exposure. Myeloperoxidase level was always lower in catla exposed at 10 and 35 °C. Therefore, both decreasing and increasing levels of temperature from the acclimation temperature affected the performance of catla. Myeloperoxidase is a microbicidal enzyme (Dalmo et al., 1997). Its elevated amount results in the destruction and elimination of invading pathogens (Yano, 1992). Higher level of myeloperoxidase is an indicator of well-developed innate immune system. Temperature tolerance of fish is dependent upon acclimation temperatures (Kasim, 2002). This phenomenon was also observed in the present study. In *Clarias batrachus*, myeloperoxidase activity reduced at temperature below and above the acclimation temperature (Singh et al., 2013). A decreasing trend of nitric oxide synthase was found with the decrease and increase of water temperature from the acclimation temperature 28 °C. NOS are a family of enzymes that catalyzes the production of cellular signalling molecule nitric oxide. This nitric oxide plays vital role in many biological processes. The inducible isoform iNOS produces large amount of nitric oxide as a defence mechanism.

The immunologically "non-permissive" temperature has been established as 4 °C in salmonids, 14 °C in carp and 22 °C in channel catfish *Ictalurus punctatus* (Bly and Clem, 1992). The innate immune response of channel catfish was influenced by low temperature (10 °C) as phagocytic activity of neutrophils reduced (Ainsworth et al., 1991). Temperature below or above the thermal limit can induce alterations to the fish immune response (Watts et al., 2001). In the present study, exposure of catla below and above acclimation temperature affected the immune system. In catla 15 °C is an immunologically "non-permissive" temperature as innate immunological parameters like, lysozyme and nitric oxide synthase activities were greatly reduced after 7 days of exposure. The young fish use innate mechanisms during the first weeks/months of their development and this may find application for the defence of farmed fish against pathogens at early age (Rombout et al., 2005). In three-spined sticklebacks (*Gasterosteus aculeatus*) the highest immune activity was found at low temperature (Dittmar et al., 2014). This may be related with the optimum temperature (18 °C) range of the species.

In present study, the heat stress increased the TBARS level in catla. Lipid peroxidation is a well-established mechanism of oxidative damage caused by ROS such as superoxide (O_2^-), OH and hydrogen peroxide (H_2O_2), and the measurement of malondialdehyde provides a convenient index of lipid peroxidation (Devasena et al., 2001). Significantly higher lipid peroxidation was found in spotted snakehead *Channa punctatus* exposed at 32 °C compared to the control group kept at 20 °C (Kaur et al., 2005). The 2.7-fold increase in malondialdehyde content was observed in juvenile seabass *Dicentrarchus labrax* after heat stress (Vinagre et al., 2012). Temperature stress caused oxidative degradation of polyunsaturated fatty acids (PUFAs) in these fishes.

The temperature exposure exerted significant modulations in expression of different heat shock proteins genes in catla. There was up-regulation of Hsp60 in muscles and hepatopancreas as the temperature reduced from acclimation one after 12 h of exposure. The expression was down regulated in both tissues after 7 days compared to the respective 12 h treatment (except in muscles at 15 °C and in hepatopancreas at 35 °C). Sharp et al. (1994) reported that in long-term

environmental stress, Hsps increased initially to protect tissues from structural damage during further exposure. In catla, higher expression of Hsp60 after 12 h exposure might help the fish to overcome the protein damage during long term exposure. Generally, higher induction of Hsps in heat-stressed animals was observed in liver suggesting that protein damage and the consequent requirement but in the present study expression of Hsp 60 was higher in muscles at the lower temperatures compared to hepatopancreas. This tissue-specific expression of Hsp 60 provided crucial information about the roles of these tissues in cold-stress. During cold acclimation, increased mitochondrial density and oxidative stress were observed in some tissues, which might increase the demand for Hsps in fishes (Orczewska et al., 2010; Teigen et al., 2015). Increased levels of Hsps in the muscles may facilitate mitochondrial biogenesis and folding of nuclear gene-encoded proteins into mitochondria. Various studies revealed that mitochondrial biogenesis is an adaptive response to cold-stress in muscles of fishes (Johnston, 1982; Tyler and Sidell, 1984; Egginton and Sidell, 1989; Orczewska et al., 2010).

In catla Hsp70 expression was up-regulated at a large extent by the heat shock. In muscles and hepatopancreas, at 35 °C the expression of Hsp70 was higher compared to control group after 7 days of exposure. Hsps protect protein substrates against conformational damage, promote the function of proteins, and prevent aggregation and formation of toxic inclusion (Tutar and Tutar, 2010). In long term exposure, Hsp 70 showed sensitivity towards increased temperature (35 °C) compared to the lower range of temperature in the present study.

The expression of Hsp90 was up-regulated in muscles at 25 and 20 °C after 12 h, and 35 °C after 7 days compared to control group. In hepatopancreas, the expression was up-regulated in 10, 15 and 35 °C treatments compared to 28 °C, after 12 h and 15, 20 and 35 °C after 7 days. Expression of Hsp90 was higher in hepatopancreas compared to muscles in catla. In heat shock treated grass carp *Ctenopharyngodon idella*, the Hsp90 expression in hepatopancreas was up-regulated (Wu et al., 2012). Elevated level of Hsp increased the stress tolerance of the fishes.

LDH gene expression has been linked with adaptation of species to environment characterize by different temperatures and dissolved oxygen concentrations (Oleksiak et al., 2002; Johns and Somero, 2004). Regulating LDH for converting between lactate and pyruvate in fish is apparently important for energy metabolism during acclimation to adverse environmental conditions. In present study, the expression of LDH-A in muscles and hepatopancreas of catla was induced in lower temperatures after 12 h exposure, whereas it was induced in warm temperature after 7 days. Seasonal acclimation to cold resulted in increased LDH in muscles of rainbow trout *Oncorhynchus mykiss* (Cordiner and Egginton, 1997) and zebra fish *Danio rerio* (McClelland et al., 2006). Regulation of LDH for converting between lactate and pyruvate in fish is apparently important for energy metabolism during stress condition (Tseng et al., 2008).

The results of the present study showed that catla had better adaptability towards higher temperatures (> 25 °C) compared to the lower temperature (< 25 °C). Temperature tolerance capability of a species is genetically controlled; it also depends on the acclimation temperature. Catla is a warm water species and was acclimated at 28 °C. Therefore, cold environmental condition (temperature < 25 °C) affected the physiology of the species and mass mortality occurred at extreme cold temperature of 10 °C. Present study showed that abiotic factor temperature played significant role in the physiology of commercially important species catla. Care should be taken for the maintenance of environmental temperature in carp culture systems.

Acknowledgements

Authors are thankful to Indian Council of Agricultural Research, ICAR, New Delhi for providing financial support in the form of NFBSFARA project (AS-2001/2010-11) to carry out the research work. Sincere thanks to the anonymous reviewers for their constructive suggestions.

References

Ahmad, T., Singh, S.P., Khangembam, B.K., Sharma, J.G., Chakrabarti, R., 2014. Food consumption and digestive enzyme activity of *Clarias batrachus* exposed to various temperatures. Aquac. Nutr. 20, 265–272.

Ainsworth, A.J., Dexiang, C., Waterstrat, P.R., Greenway, T., 1991. Effect of temperature on the immune system of channel catfish (*Ictalurus punctatus*) - I. Leucocyte distribution and phagocyte function in the anterior kidney at 10 °C. Comp. Biochem. Physiol. A Physiol. 100, 907–912.

Anderson, D.P., Siwicki, A.K., 1995. Basic haematology and serology for fish health programs. In: Shariff, M., Arthur, J.R., Subasinghe, R.P. (Eds.), Diseases in Asian Aquaculture II. Fish Health Section, Asian Fisheries Society, Manila, pp. 185.

Bendiksen, E.Å., Jobling, M., Arnesen, A.M., 2002. Feed intake of Atlantic salmon parr *Salmo salar* L. in relation to temperature and feed composition. Aquac. Res. 33, 525–532.

Bernfeld, P., 1955. Amylases, α and β. In: Enzymology. Academic Press, pp. 149–158.

Bly, J.E., Clem, L.W., 1992. Temperature and teleost immune functions. Fish Shellfish Immunol. 2, 159–171.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Brown, J.A., Pepin, P., Methven, D.A., Somerton, D.C., 1989. The feeding, growth and behaviour of juvenile cod, *Gadus morhua* L., in cold environments. J. Fish Biol. 35, 373–380.

Clark, D.S., Brown, J.A., Goddard, S.J., Moir, J., 1995. Activity and feeding behaviour of Atlantic cod (*Gadus morhua*) in sea pens. Aquaculture 131, 49–57.

Cnaani, A., 2006. Genetic perspective on stress response and disease resistance in aquaculture. Isr. J. Aquacult. Bamidgeh 58, 375–383.

Cordiner, S., Egginton, S., 1997. Effects of seasonal temperature acclimatization on muscle metabolism in rainbow trout, *Oncorhynchus mykiss*. Fish Physiol. Biochem. 16, 333–343.

Dalmo, R.A., Ingebrigtsen, K., Bøgwald, J., 1997. Non-specific defence mechanisms in fish, with particular reference to the reticuloendothelial system (RES). J. Fish Dis. 20, 241–273.

Dalvi, R.S., Pal, A.K., Tiwari, L.R., Das, T., Baruah, K., 2009. Thermal tolerance and oxygen consumption rates of the catfish *Horabagrus brachysoma* (Günther) acclimated to different temperatures. Aquaculture 295, 116–119.

Devasena, T., Lalitha, S., Padma, K., 2001. Lipid peroxidation, osmotic fragility and antioxidant status in children with acute post-streptococcal glomerulonephritis. Clin. Chim. Acta 308, 155–161.

Dittmar, J., Janssen, H., Kuske, A., Kurtz, J., Scharsack, J.P., 2014. Heat and immunity: an experimental heat wave alters immune functions in three-spined sticklebacks (*Gasterosteus aculeatus*). J. Anim. Ecol. 83, 744–757.

Egginton, S., Sidell, B.D., 1989. Thermal acclimation induces adaptive changes in subcellular structure of fish skeletal muscle. Am. J. Phys. Regul. Integr. Comp. Phys. 256, R1–R9.

Erlanger, B.F., Kokowsky, N., Cohen, W., 1961. The preparation and properties of two new chromogenic substrates of trypsin. Arch. Biochem. Biophys. 95, 271–278.

FAO Aquaculture Feed and Fertilizer Resources Information System Catla - natural food and feeding habits. http://www.fao.org/fishery/affris/species-profiles/catla/natural-food-and-feeding-habits/en/ (accessed 11.04.2017a).

FAO http://www.fao.org/fishery/culturedspecies/Catla_catla/en (accessed 11.04.2017b).

FAO, Food and Agriculture Organization, 2012. The State of World Fisheries and Aquaculture 2012. FAO Fisheries and Aquaculture Department. Food and Agriculture Organization of the United Nations, Rome, Italy.

Folkedal, O., Torgersen, T., Nilsson, J., Oppedal, F., 2010. Habituation rate and capacity of Atlantic salmon (*Salmo salar*) parr to sudden transitions from darkness to light. Aquaculture 307, 170–172.

García-Carreño, F.L., 1992. The digestive proteases of langostilla (*Pleuroncodes planipes*, decapoda): their partial characterization, and the effect of feed on their composition. Comp. Biochem. Physiol. B 103, 575–578.

Goolish, E.M., Adelman, I.R., 1984. Effects of ration size and temperature on the growth of juvenile common carp (*Cyprinus carpio* L.). Aquaculture 36, 27–35.

Johns, G.C., Somero, G.N., 2004. Evolutionary convergence in adaptation of proteins to temperature: A4-lactate dehydrogenases of Pacific damselfishes (*Chromis* spp.). Mol. Biol. Evol. 21, 314–320.

Johnston, I.A., 1982. Capillarisation, oxygen diffusion distances and mitochondrial content of carp muscles following acclimation to summer and winter temperatures. Cell Tissue Res. 222, 325–337.

Kasim, H.M., 2002. Thermal ecology: a vital prerequisite for aquaculture and related practices. In: Venkataramani, B., Sukumarn, N. (Eds.), Thermal Ecology. DAE Mumbai Publishers, Mumbai, pp. 222–234.

Kaur, M., Atif, F., Ali, M., Rehman, H., Raisuddin, S., 2005. Heat stress-induced alterations of antioxidants in the freshwater fish *Channa punctata* Bloch. J. Fish Biol. 67, 1653–1665.

Lee, D.U., Kang, Y.J., Park, M.K., Lee, Y.S., Seo, H.G., Kim, T.S., Kim, C.H., Chang, K.C., 2003. Effects of 13-alkyl-substituted berberine alkaloids on the expression of COX-II, TNF-α, iNOS, and IL-12 production in LPS-stimulated macrophages. Life Sci. 73, 1401–1412.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-

time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25, 402–408.
McClelland, G.B., Craig, P.M., Dhekney, K., Dipardo, S., 2006. Temperature- and exercise-induced gene expression and metabolic enzyme changes in skeletal muscle of adult zebrafish (*Danio rerio*). J. Physiol. 577, 739–751.
Ohkawa, H., Ohishi, N., Yagi, K., 1979. Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. Anal. Biochem. 95, 351–358.
Oleksiak, M.F., Churchill, G.A., Crawford, D.L., 2002. Variation in gene expression within and among natural populations. Nat. Genet. 32, 261–266.
Orczewska, J.I., Hartleben, G., O'Brien, K.M., 2010. The molecular basis of aerobic metabolic remodeling differs between oxidative muscle and liver of threespine sticklebacks in response to cold acclimation. Am. J. Phys. Regul. Integr. Comp. Phys. 299, R352–R364.
Pérez-Casanova, J.C., Afonso, L.O.B., Johnson, S.C., Currie, S., Gamperl, A.K., 2008. The stress and metabolic responses of juvenile Atlantic cod *Gadus morhua* L. to an acute thermal challenge. J. Fish Biol. 72, 899–916.
Pörtner, H.O., 2002. Physiological basis of temperature-dependent biogeography: trade-offs in muscle design and performance in polar ectotherms. J. Exp. Biol. 205, 2217–2230.
Quade, M.J., Roth, J.A., 1997. A rapid, direct assay to measure degranulation of bovine neutrophil primary granules. Vet. Immunol. Immunopathol. 58, 239–248.
Richter, K., Haslbeck, M., Buchner, J., 2010. The heat shock response: life on the verge of death. Mol. Cell 40, 253–266.
Rombout, J., Huttenhuis, H.B.T., Picchietti, S., Scapigliati, G., 2005. Phylogeny and ontogeny of fish leucocytes. Fish Shellfish Immunol. 19, 441–455.
Saurabh, S., Sahoo, P.K., 2008. Lysozyme: an important defence molecule of fish innate immune system. Aquac. Res. 39, 223–239.
Sharp, V.A., Miller, D., Bythell, J.C., Brown, B.E., 1994. Expression of low molecular weight HSP 70 related polypeptides from the symbiotic sea anemone *Anemonia viridis* Forskall in response to heat shock. J. Exp. Mar. Biol. Ecol. 179, 179–193.
Singh, R.K., Desai, A.S., Chavan, S.L., Khandagale, P.A., 2009. Effect of water temperature on dietary protein requirement, growth and body composition of Asian catfish, *Clarias batrachus* fry. J. Therm. Biol. 34, 8–13.
Singh, S.P., Sharma, J.G., Ahmad, T., Chakrabarti, R., 2013. Effect of water temperature on the physiological responses of Asian catfish *Clarias batrachus* (Linnaeus 1758). Asian Fish. Sci. 26, 26–38.
Siwicki, A.K., 1989. Immunostimulating influence of levamisole on nonspecific immunity in carp (*Cyprinus carpio*). Dev. Comp. Immunol. 13, 87–91.
Somero, G.N., 1998. Adaptations to cold and depth: contrasts between polar and deep-sea animals. In: Portner, H.O., Playle, R.C. (Eds.), Cold Ocean Physiology. Cambridge University Press, Cambridge, pp. 33–57.
Teigen, L.E., Orczewska, J.I., McLaughlin, J., O'Brien, K.M., 2015. Cold acclimation increases levels of some heat shock protein and sirtuin isoforms in threespine stickleback. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 188, 139–147.
Toivola, D.M., Strnad, P., Habtezion, A., Omary, M.B., 2010. Intermediate filaments take the heat as stress proteins. Trends Cell Biol. 20, 79–91.
Tseng, Y.C., Lee, J.R., Chang, J.C.H., Kuo, C.H., Lee, S.J., Hwang, P.P., 2008. Regulation of lactate dehydrogenase in tilapia (*Oreochromis mossambicus*) gills during acclimation to salinity challenge. Zool. Stud. 47, 473–480.
Tutar, L., Tutar, Y., 2010. Heat shock proteins; an overview. Curr. Pharm. Biotechnol. 11, 216–222.
Tyler, S., Sidell, B.D., 1984. Changes in mitochondrial distribution and diffusion distances in muscle of goldfish upon acclimation to warm and cold temperatures. J. Exp. Zool. 232, 1–9.
Vinagre, C., Madeira, D., Narciso, L., Cabral, H.N., Diniz, M., 2012. Effect of temperature on oxidative stress in fish: lipid peroxidation and catalase activity in the muscle of juvenile seabass, *Dicentrarchus labrax*. Ecol. Indic. 23, 274–279.
Watts, M., Munday, B., Burke, C., 2001. Immune responses of teleost fish. Aust. Vet. J. 79, 570–574.
Winkler, U.K., Stuckmann, M., 1979. Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*. J. Bacteriol. 138, 663–670.
Wu, C.X., Zhao, F.Y., Zhang, Y., Zhu, Y.J., Ma, M.S., Mao, H.L., et al., 2012. Overexpression of Hsp90 from grass carp (*Ctenopharyngodon idella*) increases thermal protection against heat stress. Fish Shellfish Immunol. 33, 42–47.
Yano, T., 1992. Assays of hemolytic complement activity. In: Stolen, J.S., Fletcher, T.C., Anderson, D.P., Kaattari, S.L., Rowley, A.F. (Eds.), Techniques in Fish Immunology. SOS Publications, USA, pp. 131–141.
Zhou, J., Wang, L., Xin, Y., Wang, W.N., He, W.Y., Wang, A.L., Liu, Y., 2010. Effect of temperature on antioxidant enzyme gene expression and stress protein response in white shrimp, *Litopenaeus vannamei*. J. Therm. Biol. 35, 284–289.

Renewable Energy 113 (2017) 503–511

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Electronic band-offsets across $Cu_2O/BaZrO_3$ heterojunction and its stable photo-electro-chemical response: First-principles theoretical analysis and experimental optimization

Dipika Sharma [a], Rishibrind Kumar Upadhyay [b], Biswarup Satpati [c], Vibha R. Satsangi [d], Rohit Shrivastav [a], Umesh V. Waghmare [e], Sahab Dass [a, *]

[a] Department of Chemistry, Dayalbagh Educational Institute, Agra, India
[b] Department of Applied Physics, Delhi Technological University, New Delhi, India
[c] Surface Physics & Material Science Division Saha Institute of Nuclear Physics Kolkata, India
[d] Department of Physics & Computer Sciences, Dayalbagh Educational Institute, Agra, India
[e] Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India

ARTICLE INFO

Article history:
Received 2 March 2017
Received in revised form
31 May 2017
Accepted 4 June 2017
Available online 7 June 2017

Keywords:
Heterojunction
Photoelectrochemical
Density functional theory
Effective mass
Band offsets

ABSTRACT

Cu_2O has been shown to be highly active as a cathode in photo-electro-chemical (PEC) reduction of water to generate hydrogen fuel using sunlight. However, Cu_2O is susceptible to photo-corrosion when exposed to an electrolyte and needs to be protected by an over-layer of a suitable material, and its performance depends sensitively on its interface with the over-layer, and hence to the method of deposition. Here, we use first-principles theoretical analysis of the electronic structure of $Cu_2O/BaZrO_3$interface, and show that valence and conduction bands are favourably aligned to absorb light and catalyse the Hydrogen Evolution Reaction. We then present experiments with spray pyrolytically deposited Cu_2O films and spin coated over-layer of $BaZrO_3$, and optimize its PEC performance with thickness of Cu_2O film. We find a maximum photocurrent density of1.25 mA/cm^2 at 0.95 V/SCE for an overall thickness of 458 nm, and demonstrate that the photocurrent remains stable over a long period of time. Demonstration of scalable and cost effective deposition of Cu_2O and $BaZrO_3$with stable PEC performance and understanding of the mechanism of charge separation across the $Cu_2O/BaZrO_3$ interface developed here should facilitate further optimization of $Cu_2O/BaZrO_3$ films for realistic PEC applications.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Generation of hydrogen by photoelectrochemical (PEC) water splitting using sun light illumination is simpler and environmentally friendlier than the catalytic reforming of hydrocarbon fuels [1,2]. Although considerable research within the past decade has been devoted into PEC water splitting, fabrication of a stable and efficient photo-electrode with Solar to Hydrogen (STH) efficiency of more than 10% required for practical applications still remains a challenge. To make this technique commercially viable, the electrode material, for example a semiconductor, must be of low-cost: it should be synthesized/fabricated in energy efficient manner with abundant elements using scalable techniques [3–5]. Use of bilayered/heterojunction system of metal oxide semiconductors with inherent merits is one of the recent strategies towards improving the performance of the photocatalyst in the recent years [6–11]. Among various semiconductors, copper oxides are attractive because they are photoactive and can be processed by industrially proven low-cost method such as spray pyrolysis, with copper being an abundant material. Cuprous oxide (Cu_2O) is an inexpensive and promising material as a photocathode in PEC for H_2 evolution through water splitting [12–14], and is also used in gas sensors [15], electronics [16] and lithium ion batteries [17]. Its electronic band gap of 2.3 eV [18,19] allows for visible light absorption, and a favourable position of its conduction band relative to redox potential of hydrogen evolution reaction provides a potential of about 0.7 V as driving force for water reduction [20]. However, a major technical issue associated with Cu_2O is its relatively short electron diffusion length (10–100 nm) [21] and high recombination rate, i.e. electrons and holes excited by incident light

* Corresponding author.
E-mail address: drsahabdas@gmail.com (S. Dass).

http://dx.doi.org/10.1016/j.renene.2017.06.022
0960-1481/© 2017 Elsevier Ltd. All rights reserved.

cannot be transferred away efficiently as they recombine readily [22]. In addition, energy of the valence band of Cu_2O is more positive than the water oxidation potential. Thus, its instability and susceptibility to photocorrosion in aqueous solution during illumination pose serious limitation to its cyclability [23], and a coating of another semiconductor is required to stabilize its surface while maintaining efficient hydrogen generation. In recent years Cu_2O has been explored as a photocatalyst with improved photoresponse with costly deposition methods like atomic layer deposition, chemical vapour deposition, and with coatings of expensive precious or noble metals like Pt and Au [23,24]. It is desirable to explore the performance of Cu_2O based photocathode fabricated with cost-effective methods of deposition such as spray pyrolysis, spin coating. Charge separation across the interface between Cu_2O and the coating material is key to the PEC performance. Electronic structure and band-offsets across the interface need to be determined to understand the mechanism of charge separation, and optimize the PEC properties of Cu_2O based photocathodes.

$BaZrO_3$ is a stable photocatalyst with perovskite cubic structure, and its valence and conduction band edges aligned favourably with respect to redox potential of water for PEC water splitting [25,26]. On the other hand $BaZrO_3$ is a wide band gap material, absorb in UV region which limits its efficiency towards photoelectrochemical splitting of water. Motivation of present work lies in the fact that combining Cu_2O with $BaZrO_3$ will only have a slight lattice mismatch (lattice constants of 4.26 Å and 4.22 Å respectively [27,28], as a single photoelectrode, and may avoid recombination of charge carriers and provide better stability for efficient PEC water splitting hydrogen generation reaction.

With this idea, we present here first-principles theoretical analysis of band offsets at $Cu_2O/BaZrO_3$ interface, to understand the charge carrier separation at the $Cu_2O/BaZrO_3$ interface. We then present experimental validation with measurement of the photo response of $Cu_2O/BaZrO_3$ thin films with varying thickness of Cu_2O for the first time, and determine the thickness for optimal performance. The $Cu_2O/BaZrO_3$ photocathode was fabricated using cost effective and scalable methods of deposition, and serves as a low cost standard electrode for photo-electrochemical H_2 production in real applications.

2. Experimental section

2.1. Synthesis of Cu_2O thin film

Nanostructured Cu_2O thin films with varying thickness were deposited on ITO (In_2O_3: Sn) glass substrate using spray pyrolysis (Holmarc, India) technique. Precursor solution comprising of 0.04M Copper (II) acetate monohydrate ($Cu(CH_3COO)_2 \cdot H_2O$ and 0.04M Dextrose dissolved in water were used as starting compounds. In addition 20 vol% of 2-propanol ($(CH_3)_2CHOH$) was added to the above described aqueous solution. Detailed methodology has been reported elsewhere [29]. To optimize the thickness of Cu_2O films 10, 20, 40 and 60 s spray periods were tried for the deposition. One third length of ITO substrate was initially covered by transparent tape to establish the electrical contact to convert them into electrodes.

2.2. Fabrication of $Cu_2O/BaZrO_3$ heterojunction electrodes

$Cu_2O/BaZrO_3$ heterojunction electrodes were fabricated by overlayering of $BaZrO_3$ on to the predeposited Cu_2O using sol gel spin coating method. For the synthesis of $BaZrO_3$ thin film, a gel was obtained by mixing zirconium butaoxide and methoxy ethanol, and then the solution was stirred at 70 °C for 30 min and cooled down to room temperature. Barium acetate in acetic acid was dissolved separately, and both the solutions were mixed while stirring for 30 min on a magnetic stirrer, resulting in a white colloidal gel, which was spin-coated onto predeposited Cu_2O thin film at 3000 rpm. After coating, the film was sintered for 2 h at 600 °C under vacuum ($p \sim 10^{-5}$ torr). The final product exhibited a form of typical white coloured stable film [30]. A schematic diagram of irradiated the $Cu_2O/BaZrO_3$ heterojunction photoelectrode is shown in Fig. S1. Thicknesses and other sample details have been summarized in Table 1. Films were converted into electrodes using copper wire, silver paste and epoxy (Hysol, Singapore) for their use as photoelectrode in PEC cell. The effective area of the photoelectrode available for illumination was 1.0 cm^2.

3. Results and discussion

3.1. Theoretical results

3.1.1. Structural model of $Cu_2O/BaZrO_3$ interface

We modelled $Cu_2O/BaZrO_3$ interface with a periodic supercell containing a vacuum and a slab consisting of four unit cells of Cu_2O stacked along (001) direction, interfacing across its (001) atomic plane 4 unit cells of $BaZrO_3$ stacked along (001) direction (see Fig. 1 (a)). The slab in the supercell is thick enough to maintain low interaction of the $Cu_2O/BaZrO_3$ interface with the two surfaces of the slab. As the atomic arrangement in the Cu_2O part of the slab is polar along the c-axis, and a thick layer of vacuum is used to weaken the effects of polarity and interaction between periodic images associated with the supercell. The interface is formed between BaO (001) plane of $BaZrO_3$ and (001) plane containing O atoms of Cu_2O. As the lattice constants of the two oxides here are within a percent of each other, we used average lattice constant for the in-plane periodicity of the interface (and the supercell) [31,32].

3.1.2. Band offsets

The structural discontinuity across the interface results in modification of the electronic states at the interface that smoothly transforms valence and conduction bands on one side to those on the other. An overall jump in the energy of valence (or conduction) band across the interface is called a band-offset. These band offsets govern the charge transport or separation across the interface and are key to efficiency of conversion of energy of photo-excited carriers to other energy forms. In theoretical analysis of a model interface, the band offsets (BO) is contributed by two terms [33,34]:

$$BO = \Delta E_{band}(v \text{ or } c) + \Delta V_{av}, \quad (1)$$

where the first contribution, $\Delta E_{band}(v \text{ or } c)$ is the band structure contribution, and defined as the difference between the valence and conduction band energies of bulk materials relative to their average electrostatic potentials. The second term, ΔV_{av}, is the discontinuity or the jump in average electrostatic potential across the heterojunction interface [35], and is obtained using coarse-graining scheme of interfacial planar and macroscopic averages of

Table 1
Thicknesses with other details of samples.

Film thickness (nm) with an error of ±10%.			
Sample ID	Cu_2O	$BaZrO_3$	overall thickness
Cu_2O	150	–	150
$BaZrO_3$	–	148	148
$Cu_2O/BaZrO_3$ (A)	152	148	300
$Cu_2O/BaZrO_3$ (B)	253	148	401
$Cu_2O/BaZrO_3$ (C)	310	148	458
$Cu_2O/BaZrO_3$ (D)	452	148	600

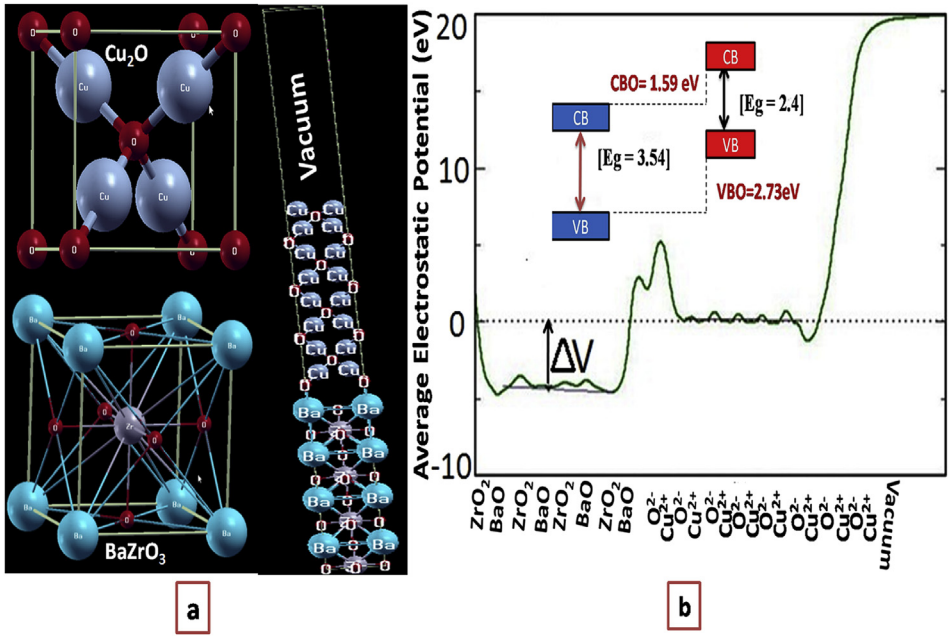


Fig. 1. (a) Unit cell of Cu_2O and $BaZrO_3$ crystals and a periodic supercell used to model their heterojunction. (b) Macroscopically averaged electrostatic potential along the direction perpendicular to $Cu_2O/BaZrO_3$interface, along with a schematic representation of their valence and conduction bands revealing the band offsets at the $Cu_2O/BaZrO_3$ interface.

the electrostatic potential. It includes the effects of charge transfer, dipoles and other structural details at the interface, and is estimated from a high accuracy calculation of the model interface.

Macroscopically averaged electrostatic potential along the direction perpendicular to $Cu_2O/BaZrO_3$ interface, along with a schematic representation of their valence and conduction bands revealing the band offsets at the $Cu_2O/BaZrO_3$ interface is shown in Fig. 1(b). Average electrostatic potential for $BaZrO_3$ and Cu_2O are also indicated and difference in these potential values at the interface resulted in the macroscopically averaged electrostatic potential along the direction perpendicular to $Cu_2O/BaZrO_3$ interface which is denoted by ΔV_{av}. While the energies of valence band edges are accurately captured in simple LDA or GGA-based DFT calculations, band-gaps are typically underestimated and need to be estimated with self-interaction corrected or hybrid functional-based calculations. Since the number of atoms in our supercell and model interface is prohibitively large for hybrid functional or self-interaction corrected calculations, we used (a) valence band-offsets from DFT calculations, and (b) experimental band-gaps of the bulk Cu_2O and $BaZrO_3$ to determine the conduction band-offsets from the former. Results of our analysis of the band offsets show that (001) $Cu_2O/BaZrO_3$ interface is of type II,*i.e.* each of the valence and conduction band edges of $BaZrO_3$are below the corresponding Cu_2O counterparts. This band-alignment allows ready transfer of electrons from the conduction band of Cu_2O to that of $BaZrO_3$, and that of holes from $BaZrO_3$ to Cu_2O. Thus, it favours efficient separation of photo-excited charge carriers and hence has high photo-electro-chemical activity of the $Cu_2O/BaZrO_3$ heterojunction (see Fig. 7 (a)).

Furthermore, mobility of charge carriers (electrons and holes) is an important contributing factor to their role in the photo-electrochemical activity. The mobility (μ) depends inversely on the effective mass of carriers, we estimated the effective masses of electrons(m_e) and holes (m_h) of bulk $BaZrO_3$ and Cu_2O from their band dispersion at the conduction band minimum and valence band maximum, using parabolic fits the band energies: $E = \frac{\hbar^2 k^2}{2m_e m_0}$, where m_0 denotes the electron rest mass [36]. Our estimates of effective masses of electrons and holes in bulk Cu_2O and $BaZrO_3$ (details of our calculations and results for electronic structure and density of states (DOS) of bulk Cu_2O and $BaZrO_3$ are given in supporting information) are listed in Table 2 [37]. While electronic effective masses of the two oxides are comparable with a conduction band-offset of 1.59 eV giving good interfacial mobility, effective masses of holes on the two sides are markedly different (see Table 2). Since the valence band offs-set is relatively larger (of 2.73 eV), we expect the interfacial mobility of holes to be reasonable as well, due to the electric field across the interface.

Calculated values of effective mass and band offsets suggest improved separation of Charge carriers at the interface [38].

3.2. Experimental results

3.2.1. XRD and SEM analysis

The XRD patterns of pristine $BaZrO_3$, Cu_2O and $Cu_2O/BaZrO_3$ heterojunction samples (A-D) (shown in Fig. 2(a)) reveal (a) the peaks at $2\theta = 36.53, 42.43$ indexed to (111), (200) plane of cubic cuprous oxide phase (JCPDS file no. 065-3288) [27], and (b) peaks at 2θ values 30.27, 43.02, 50.37 corresponds to(110), (111), (211) planes of cubic phase of barium zirconate ((JCPDS file no. 006-0399) [28]. The absence of any unidentified peaks in the case of heterojunction samples indicates that no mixed oxide phase forms. However, we note that the observed XRD peaks are broad with significant background. Such broad XRD peaks, with high background, have also been reported in case of very small sized Cu_2O nanocrystals [39].

The nanostructured surface morphology of photoelectrodes

Table 2
Components of the effective mass tensors (in unit of m_0, $m_0 = 9.1 \times 10^{-31}$Kg).

	Cu_2O	$BaZrO_3$
CBM at Γ point Γ-R (111) Direction	0.676	0.99
VBM at Γ point Γ-R (111) Direction	3.80	–
VBM at M point M-X (110) Direction	–	1.37

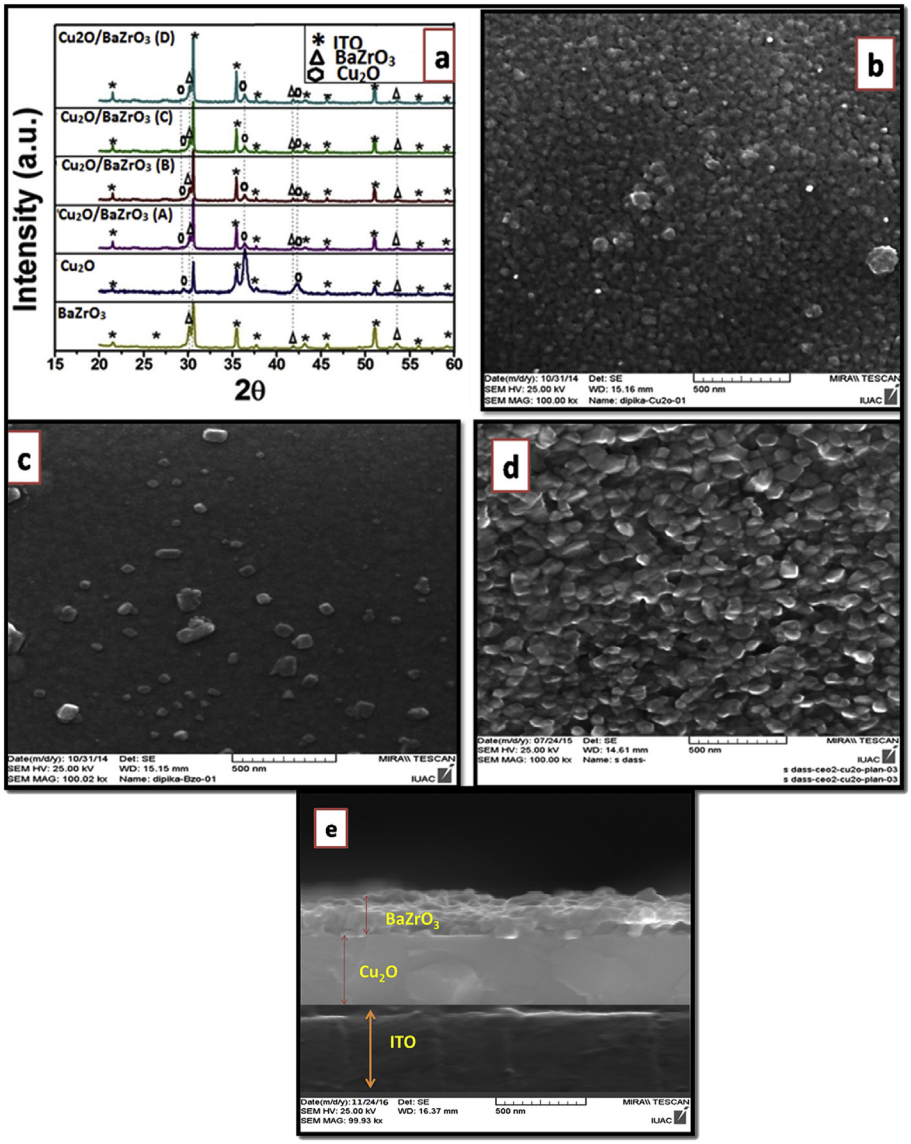


Fig. 2. (a) X-ray diffraction patterns for pristine Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$heterojunctions (A–D). SEM images of (b) pristine Cu_2O (c) the pristine $BaZrO_3$ (d) $Cu_2O/BaZrO_3$ heterojunction sample 'C' (e) Cross Sectional SEM image for $Cu_2O/BaZrO_3$ heterojunction sample 'C'.

plays an important role in their PEC performance. In particular, several studies have demonstrated that organized nanostructures can improve the minority-carrier transport and reduce the rate and extent of electron/hole recombination [40]. The SEM images of pristine Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$ heterojunction sample 'C' are displayed in Fig. 2(b–d). An SEM image of $Cu_2O/BaZrO_3$ shows that the underlying Cu_2O film is madeup of uniform nano-sized particles covered by $BaZrO_3$ porous layer (Fig. 2 (d)). We have assessed the purity and structural quality of these films using EDAX. Cross sectional SEM image (side view) of $Cu_2O/BaZrO_3$ heterojunction sample'C' indicate the deposition of two layers. Thickness for $BaZrO_3$ and Cu_2O layers as calculated is 160 nm and 330 nm respectively and are in good agreement with the values obtained from surface profilometer measurements (Fig. 2(e)). AFM images (see the relevant details given in the supporting information).

3.2.2. XPS analysis

XPS data for analysing the surface chemical composition of $Cu_2O/BaZrO_3$ heterojunction sample 'C' were collected. As shown in Fig. 3 (A), signals corresponding to Cu, Ba, Zr, and O were noted and a peak for C 1s at 284.8 eV is ascribed to adventitious carbon from the XPS instrument [8]. In the survey scan a low intensity peak for Cu is obtained as compared to Ba and O which may be due to presence of Cu_2O at the inner surface of the $Cu_2O/BaZrO_3$ heterojunction. As XPS is surface sensitive technique only Ba, Zr and O are appears in high resolution XPS for $Cu_2O/BaZrO_3$ heterojunction thin film Fig. 3 (B–D). Peaks at 538.7 eV is for bonded O 1s and a small intensity board peak at 531.68 eV is for lattice oxygen (small intensity may be due to vacuum annealing) [41,42]. Whereas the Zr $3d_{3/2}$ and $3d_{5/2}$ peaks are situated at 183.23 eV and 180.87 eV respectively. Data obtained are consistent with the +2 oxidation states of Ba and +4 for Zr, respectively, which are expected to $BaZrO_3$ formation. The XPS results further confirmed the coexistence of Cu_2O and $BaZrO_3$ in the $Cu_2O/BaZrO_3$ heterojunction [43,44], agreeing well with the XRD results.

3.2.3. TEM analysis

For structural analysis, transmission electron microscopy (TEM)

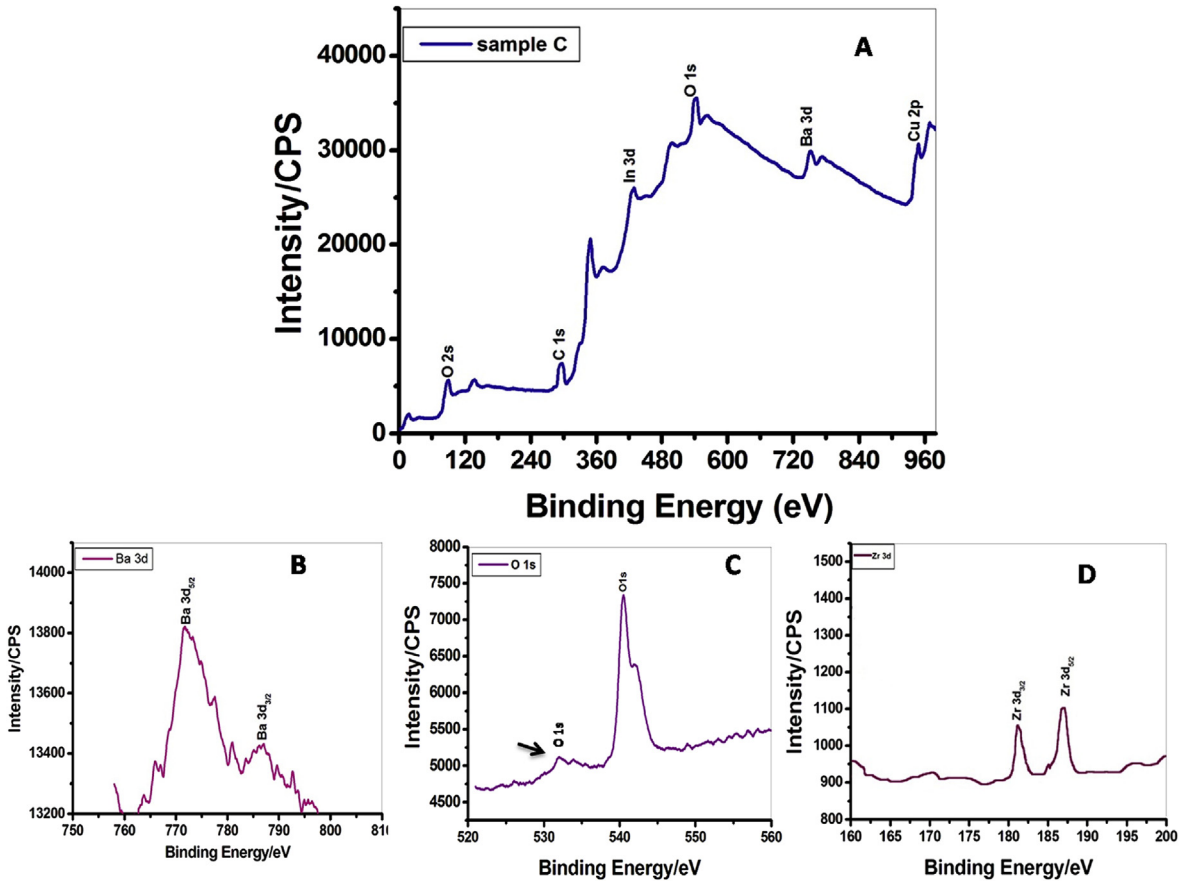


Fig. 3. XPS spectra of (A) full-spectrum scan, (B) Cu 2p, (C) Ba 3d, and (D) O 1s (E) Zr 3d for $Cu_2O/BaZrO_3$ heterojunction sample 'C'.

was performed in cross-sectional geometry for $Cu_2O/BaZrO_3$ heterojunction sample 'C' shown in Fig. 4. To investigate the chemical composition of the heterojunction and detailed distribution of atomic contents in the materials, elemental mapping was performed using STEM-HAADF-EDX technique (Fig. 4 a–b) which shows the presence of Cu, O, Ba, and Zr element with no other impurity. Fourier filtered image from the interface, which shows some evidence of epitaxial alignment at the interface (Fig. 4 c). In

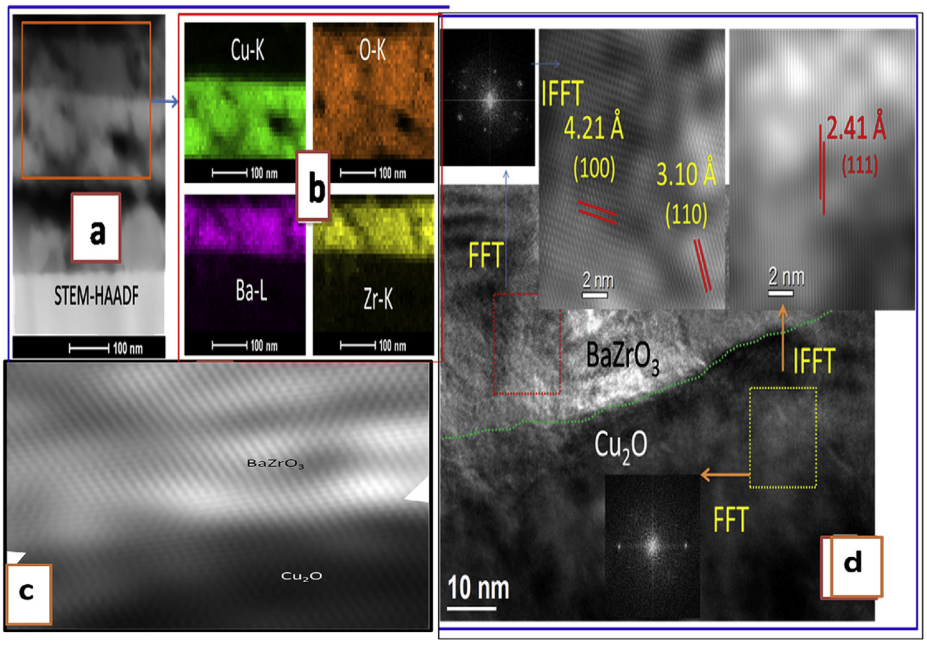


Fig. 4. (a) HAADF-STEM image of $Cu_2O/BaZrO_3$ heterojunction sample 'C', (b) corresponding EDX elemental mapping of Cu, O, Ba, and Zr, (c) Fourier filtered image from the interface (d) HRTEM image of the interface and corresponding FFT and IFFT.

the high-resolution (HR) cross-sectional TEM (XTEM) image of the Cu_2O and $BaZrO_3$ interface is shown in Fig. 4 d. Fast Fourier transform (FFT) and corresponding inverse FFT (IFFT) from two different regions marked by yellow and red dotted box (namely, Cu_2O and $BaZrO_3$) show crystalline quality of these layers. The inter planer spacing (d-spacing) estimated from the HRTEM images can be sequentially referred to cubic Cu_2O and $BaZrO_3$. Therefore, both HRTEM and EDX characterizations indicate the formation of a clean $Cu_2O/BaZrO_3$ heterojunction on the ITO substrate.

3.2.4. Absorption spectra

UV–Vis absorption spectra of pristine Cu_2O,$BaZrO_3$ and all $Cu_2O/BaZrO_3$ heterojunction (A-D) samples with 150 nm, 148 nm, 300 nm, 401 nm, 458 nm, 600 nm overall thickness respectively, are shown in Fig. 5. The absorption edges of pristine $BaZrO_3$ and Cu_2O are at 350 nm, 525 nm, exhibiting the absorption in UV and Visible regions respectively. Thin films with a $Cu_2O/BaZrO_3$heterojunction do absorb in the UV region, and their absorption extends up to the visible region with a red shift from 350 nm to 550 nm in absorption edge, as the thickness of Cu_2O increases. This clearly indicates the influence of Cu_2O on the absorption by these thin films [45]. The observed broad absorption from 450 to 550 nm (exhibited by films with the $Cu_2O/BaZrO_3$ heterojunction) suggests that photons can be simultaneously absorbed by both $BaZrO_3$ and Cu_2O, and the interfacial structure developed here has significantly enhanced light absorption ability.

3.2.5. Photoelectrochemical performance

We investigated photoelectrochemical performance of the samples with heterojunction by employing them as photocathodes in the photoelectrochemical cell designed for water splitting reaction. For comparison, pristine Cu_2O and $BaZrO_3$ thin films were also tested under the same setup and conditions. The photocurrent density versus the applied voltage of all the samples (see Fig. 6 (a)) shows negligible current in the dark conditions. Upon illumination of light, the sample with $Cu_2O/BaZrO_3$ heterojunction shows p-type photocurrent characteristics which increases with increase in the cathodic bias. Maximum photocurrent density of 1.25 mA/cm^2 is exhibited $Cu_2O/BaZrO_3$ heterojunction sample 'C, which is much higher than that obtained value of 0.1 mAcm^{-2} generated by pristine Cu_2O used as a cathode at 0.95 V/SCE. The reduction in photocurrent at higher thickness of Cu_2O (452 nm) is likely due to the excessive amount of

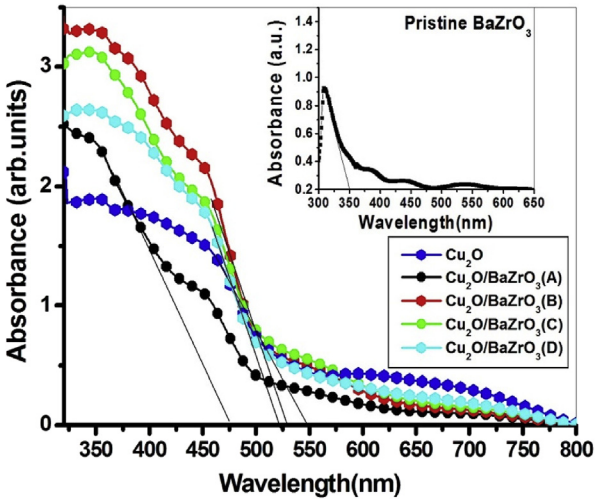


Fig. 5. UV–visible absorption spectra of pristine Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$heterojunction thin film samples (A–D).

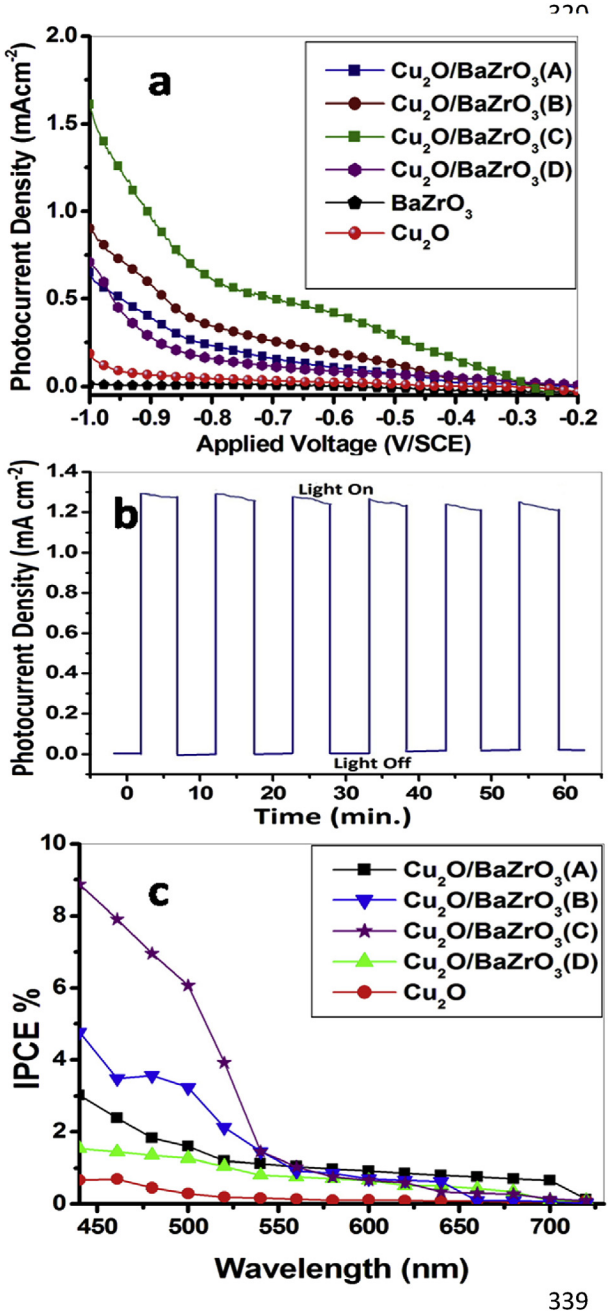


Fig. 6. (a) Photocurrent density of pristine Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$heterojunction samples ((A–D) as a function of applied voltage, (b) PEC Photocurrent density of $Cu_2O/BaZrO_3$heterojunction sample'C' at 0.95 V/SCEunder extended–chopped illumination showing its remarkable stability. (c) IPCE performance of pristine Cu_2O, and $Cu_2O/BaZrO_3$ heterojunction samples (A–D).

Cu_2O which act as recombination centers: photogenerated e^-/h^+ pairs can recombine before they can reach to the ITO layer [46]. The improved photoresponse of $Cu_2O/BaZrO_3$ heterojunction can be explained with the following arguments: i) the interface allows better charge separation across the interface due the electric field generated by the band-offsets demonstrated theoretically, ii) nanostructures provide an effective percolation pathway for H^+ and H_2O and more reactive sites with their high surface-to-volume ratio, and (iii) more effective harvesting of the photons. Other possible factors contributing to the marked increase in photoresponse of these samples include decrease in resistivity [47].

3.2.6. Photostability

Samples with a $Cu_2O/BaZrO_3$ heterojunction were found to be stable under applied potential of 0.95 V/SCE for 60 min illumination under UV–visible light source in 0.1M NaOH aqueous electrolyte. No loss in the photocurrent generation implies a remarkable stability due to the overlayer of stable $BaZrO_3$ thin film onto Cu_2O thin film. Observations made under chopped illumination clearly show role of light in photocurrent generation by $Cu_2O/BaZrO_3$ heterojunction sample 'C' (Fig. 6 (b)) highlight and confirm the stability.

3.2.7. Flatband potential

The flat-band potential (V_{fb}) of a semiconductor plays an important role in deciding its PEC activity and can be calculated using the Mott–Schottky equation (given below), which relates semiconductor-electrolyte interfacial capacitance (C) to applied voltage (V_{app}).

$$\frac{1}{C^2} = \left(\frac{2}{q\varepsilon_0\varepsilon_s A^2 N_D}\right)\left(V_{app} - V_{fb} - \frac{k_B T}{q}\right)$$

where C is the space-charge capacitance, A (cm [2]) is the active geometric area, N_D (cm^{-3}) is the donor density, q is the electronic charge (1.60210^{-19} C), ε_s is the relative permittivity of semiconductor, ε_0 is the vacuum permittivity (8.854×10^{-12} F m^{-1}), V_{app} (V vs. SCE) is the applied potential, V_{fb} (V vs. SCE) is the flat-band potential, k_B is the Boltzmann constant (1.381×10^{-23} J K^{-1}) and T is the temperature (298 K). Capacitance (C) at the semiconductor/electrolyte junction was measured using an LCR meter (Agilent technology, Model 4263B) at varying electrode potentials of the AC signal frequency of 1 kHz, and the Mott-Schottky behaviour ($1/C^2$ vs applied potential) was obtained and analyzed for all the samples. Flat band potentials of Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$ heterojunction sample 'C' were calculated from the slopes of their Mott-Schottky curves as shown in Fig. 7. Extrapolation of $1/C^2$ vs. V plots to the voltage axis gives the magnitude of flat band potential (within k_BT). Larger positive value of flatband potential results in the improved charge separation. It was found that V_{fb} increases from 0.15 V/SCE for pristine Cu_2O to 0.61 V/SCE for $Cu_2O/BaZrO_3$ heterojunction sample 'C'. The highest value of the flatband potential obtained for sample 'C' is consistent with its highest photocurrent density. Charge transfer process occurring at the interface between photoelectrodes and the electrolyte was also investigated using electrochemical impedance spectroscopy (EIS) (from Nyquist plots) measurements under illumination. Fig. 7(b) shows Nyquist plots of pristine Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$ heterojunction sample 'C'. The plot shows that semicircle of the plot is getting smaller for $Cu_2O/BaZrO_3$ (C) as compared to pristine Cu_2O, $BaZrO_3$ confirming effective enhancement of charge transfer [48].

3.2.8. Charge separation mechanism

The charge transfer process between $BaZrO_3$ and Cu_2O, and the energy level diagram of $Cu_2O/BaZrO_3$ heterojunction are illustrated in Fig. 8(a–b). A p–p heterojunction between Cu_2O and $BaZrO_3$ was formed after the deposition of $BaZrO_3$ on the Cu_2O thin films. Relative band energies of Cu_2O and $BaZrO_3$ offer the band potential gradient that favours the photogenerated electron-hole separation at the $Cu_2O/BaZrO_3$ interface. The band edge positions of $BaZrO_3$ and Cu_2O were calculated using equation $E_{VB} = X - E_e + 0.5E_g$ and $E_{CB} = E_{VB} - E_g$ Refs. [49,50], where E_e is energy of free electron on the hydrogen scale (−4.5eV), $E_g$is the band gap and X is the absolute electronegativity of the semiconductor (geometric mean of electronegativity of constituents atoms). The difference in the valence and conduction bands of $BaZrO_3$ and Cu_2O could lead to a band offset, which agrees well with results of our theoretical analysis presented earlier in this section. UV–Visible light irradiation on

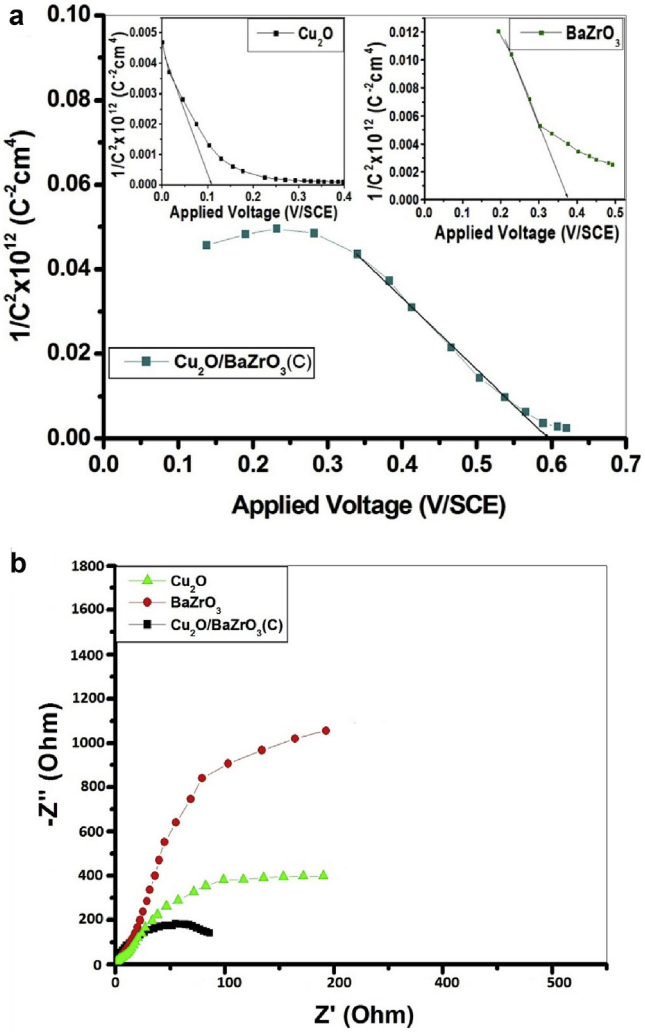


Fig. 7. (a) Mott-Schottky curve of $Cu_2O/BaZrO_3$ heterojunction sample 'C' with the Mott-Schottky curves of pristine Cu_2O, $BaZrO_3$ in the insets. (b) Nyquist plots of pristine Cu_2O, $BaZrO_3$ and $Cu_2O/BaZrO_3$ heterojunction sample 'C' under simulated sunlight irradiation.

heterojunction results in band gap excitation of the $BaZrO_3$ and Cu_2O thereby creating an electron–hole pair. Due to more positive conduction band potential of $BaZrO_3$ than that of Cu_2O, the photogenerated electrons from Cu_2O can readily transfer to the conduction band of $BaZrO_3$. On the other hand, because of the higher valence band potential of $BaZrO_3$, than Cu_2O the photogenerated holes in the $BaZrO_3$ are able to easily transfer to the valence band of Cu_2O preventing the photocorrosion of Cu_2O. Thus, the charge carriers can be efficiently separated and transferred. The photo-induced hole in the valence band of Cu_2O is transferred to the external circuit and consumed by participating in water oxidation at counter electrode whereas accumulated electrons in the conduction band of $BaZrO_3$ reduce the water to form hydrogen. The overall process of generation and separation of electron-hole pairs increases due to presence of the additional $BaZrO_3$ layer. This effect leads to the enhancement of photocurrent densities [51].

3.2.9. Estimation of efficiency

The incident photon to electron conversion efficiency (IPCE) of the samples was measured using electrochemical workstation (Zahner, PP211, CIMPS-pcs, Germany) with wideband white LED

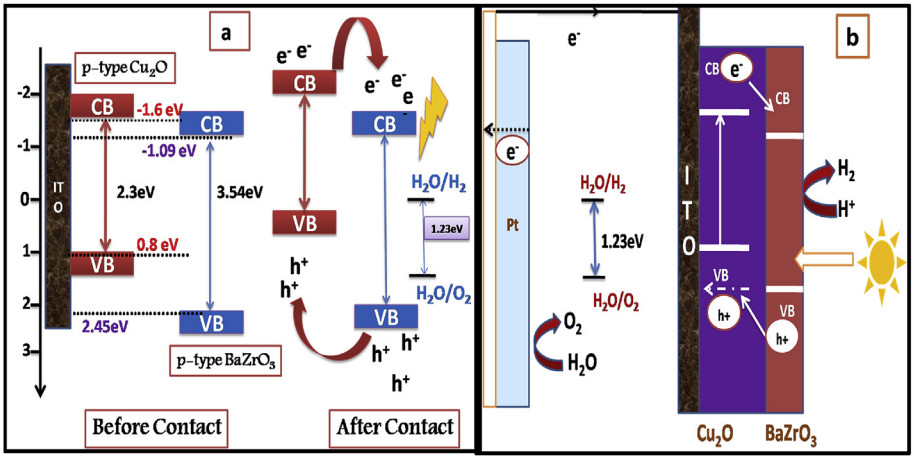


Fig. 8. (a) Electronic energy bands of Cu_2O and $BaZrO_3$ before and after the formation of *p-p* junction. (b) A schematic illustration of solar water splitting in the two electrodes PEC system, using $Cu_2O/BaZrO_3$photocathode and the Pt mesh as a photoanode.

tuneable light source (output intensity = 1.02 W/m^{-2}) and estimated with

$$IPCE = \frac{1240I}{\lambda J_{Light}},$$

where J is the measured photocurrent density, λ is the wavelength of the incident light and I is the measured irradiance with light [52]. The IPCE values of all the samples were obtained in the wavelength range of 440–730 nm, as shown in Fig. 6 (c). $Cu_2O/BaZrO_3$ heterojunction sample 'C' showed the highest IPCE of over 8.8% at 480 nm wavelength, which gradually decreased to almost zero at 720 nm. A significant increase in the IPCE indicates that the sample with Cu_2O:$BaZrO_3$ heterojunction is definitely superior in harvesting the incident photons and inducing water splitting [53]. Multiple studies on various other nanosystems indicate that, this improvement could be attributed to a combination of better incident photon absorption and improved charge separation due to the well-ordered nanostructures. In addition, it has been reported that the electron transfer from Cu_2O to $BaZrO_3$ results in effective electron-hole separation, which could further enhance the IPCE [54]. Applied bias photon-to-current efficiency (ABPE) was also calculated for all the samples and values are supporting the PEC results (See Table S2).

4. Conclusions

From first-principles theoretical analysis of $BaZrO_3$:Cu_2O interface, we showed that the conduction and valence bands of $BaZrO_3$ are offset by1.59 eV and 2.73 eV below the conduction and valence bands of Cu_2O respectively. Based on this and the fact that crystal structures of the two oxides are almost lattice matched, we predicted that a coherent interface between $BaZrO_3$ and Cu_2O is quite favourable to electron-hole separation and PEC properties of the latter for hydrogen evolution reaction of the latter. $BaZrO_3$ is thus a good candidate for an over-layer on Cu_2O to improve its stability against photo-corrosion while maintaining its PEC performance. We then developed thin films of controlled thickness of Cu_2O using spray pyrolysis and an over-layer of $BaZrO_3$ using spin-coating, and established the structural purity and quality of the $Cu_2O/BaZrO_3$ interface with thorough characterization diffraction, spectroscopic techniques and microscopy. We demonstrated excellent absorption in UV and visible range by these films with a remarkably stable PEC performance in splitting water to generate hydrogen, achieving a maximum IPCE efficiency of 8.8% exhibited by $Cu_2O/BaZrO_3$ films of overall thickness 458 nm. Abundance of materials used here, cost-effective deposition techniques used in fabricating these films, with stable and reasonable PEC performance, make them promising photo-electrodes for use in realistic applications to generate hydrogen from solar water splitting.

Acknowledgements

Authors gratefully acknowledge financial support received from the Department of Science & Technology, New Delhi, India, vide project no. SR/NM/NS-147/2010. Dipika Sharma is thankful to DST for providing Inspire Fellowship.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.renene.2017.06.022.

References

[1] A. Fujishima, A.T. Rao, D. Tryk, Titanium dioxide photocatalysis, photochemistry reviews, J. Photochem. Photobiol. C 1 (2000) 1–2.
[2] T. Bak, J. Nowotny, M. Rekas, C.C. Sorrell, Photo-electrochemical hydrogen generation from water using solar energy: materials-related aspects, Int. J. Hydrogen Energy 27 (2002) 91–1022.
[3] O. Khaselev, J.A. Turner, A monolithic photovoltaic-photoelectrochemical device for hydrogen production via water splitting, Science 280 (1998) 425–427.
[4] S.E. Aharon, Heller, Efficient p-InP (Rh-H alloy) and p-InP(Re–H alloy) hydrogen evolving photocathodes, J. Electrochem. Soc. 129 (1982) 865–866.
[5] B. Marsen, B. Cole, E.L. Miller, Photoelectrolysis of water using thin copper gallium diselenide electrodes, Sol. Energy Mater. Sol. Cells 92 (2008) 1054–1058.
[6] K. Guoa, Z. Liua, C. Zhoua, J. Hana, Y. Zhaob, Y. Lia, T. Cuia, B. Wanga, J. Zhang, Fabrication of TiO_2 nano-branched arrays/Cu2S composite structure and its photoelectric performance, Appl. Catal. B Environ. 154–155 (2014) 27–35.
[7] X. Chang, T. Wang, P. Zhang, Y. Wei, J. Zhao, J. Gong, Stable aqueous photoelectrochemical CO_2 reduction by a Cu_2O dark cathode with improved selectivity for carbonaceous products, Angew. Chem. Int. Ed. 55 (2016) 8840–8845.
[8] X. Chang, T. Wang, P. Zhang, J. Zhang, A. Li, J. Gong, Enhanced surface reaction kinetics and charge separation of p–n heterojunction Co3O4/BiVO4 photoanodes, J. Am. Chem. Soc. 137 (2015) 8356–8359.
[9] C. Li, A. Li, Z. Luo, J. Zhang, X. Chang, Z. Huang, T. Wang, J. Gong, Surviving high-temperature calcination: ZrO_2-induced hematite nanotubes for photoelectrochemical water oxidation, Angew. Chem. Int. Ed. 56 (2017) 4150–4155.
[10] D. Sharma, A. Verma, V.R. Satsangi, R. Shrivastav, S. Dass, Nanostructured $SrTiO_3$ thin films sensitized by Cu_2O for photoelectrochemical hydrogen generation, Int. J. Hydrogen Energy 39 (2014) 4289–4297.
[11] D. Sharma, V.R. Satsangi, R. Shrivastav, S. Dass, Understanding the photoelectrochemical properties of nanostructured CeO_2/Cu_2O heterojunction

photoanode for efficient photoelectrochemical water splitting, Int. J. Hydrogen Energy 41 (2016) 18339–18350.
[12] J. Morales, L. Sanchez, S. Bijani, L. Martınez, M. Gabas, Electrodeposition of Cu_2O: an excellent method for obtaining films of controlled morphology and good performance in Li-ion batteries, Electrochem. Solid-State Lett. 8 (2005) A159–A162.
[13] Y.H. Lee, I.C. Leu, C.L. Liao, S.T. Chang, M.T. Wu, J.H. Yen, K.Z. Fung, Fabrication and characterization of Cu_2O nanorod arrays and their electrochemical performance in Li-ion batteries, Electrochem. Solid-State Lett. 9 (2006) A207–A210.
[14] Y.D. Li, J.T. Zhang, J.F. Liu, Q. Peng, X. Wang, Nearly monodisperse Cu_2O and CuO nanospheres: preparation and applications for sensitive gas sensors, Chem. Mater. 18 (2006) 867–871.
[15] J.Y. Ho, M.H. Huang, Synthesis of submicrometer-sized Cu_2O crystals with morphological evolution from cubic to hexapod structures and their comparative photocatalytic activity, J. Phys. Chem. C 113 (2009) 14159–14164.
[16] J.M. Tarascon, P. Poizot, S. Laruelle, S. Grugeon, L. Dupont, Nano-sized transition- metal oxides as negative-electrode materials for lithium-ion batteries, Nature 407 (2000) 496–499.
[17] W. Siripala, J.R.P. Jayakody, Observation of n-type photoconductivity in electrodeposited copper oxide film electrodes in a photoelectrochemical cell, Sol. Energy Mater 14 (1986) 23–27.
[18] A.E. Rakhshani, Preparation, characteristics and photovoltaic properties of cuprous oxide - a review, Soild State Electron. 29 (1986) 7–17.
[19] W.L. Siripala, D.R.D. Perera, K.T.L. De Silva, J.K.D.S. Jayanetti, I.M. Harmadasa, Study of annealing effects of cuprous oxide grown by electrodeposition technique, Sol. Energy Mater. Sol. Cells 44 (1996) 251–260.
[20] P.E. Jongh, D. de, Vanmaekelbergh, J.J. Kelly, Cu_2O: electrodeposition and characterization, Chem. Mater 11 (1999) 3512–3517.
[21] P. E. de Jongh, D. Vanmaekelbergh, J.J. Kelly, Photoelectrochemistry of electrodeposited Cu_2O, Electrochem. Soc. 147 (2000) 486–489.
[22] S.J. Wang, S.F. Yin, L. Li, B.Q. Xu, C.F. Ng, C.T. Au, Investigation on modification of Ru/CNTscatalyst for the generation of COx-free hydrogen from ammonia, Appl. Catal. B 52 (2004) 287–299.
[23] A. Paracchino, V. Laporte, K. Sivula, M. Grätzel, E. Thimsen, Highly active oxide photocathode for photoelectrochemical water reduction, Nat. Mater. (2011) 456–461.
[24] Z. Zhang, P. Wang, Cu_2O/CuO photocathode with improved stability for photoelectrochemical water reduction, J. Mater. Chem. 22 (2012) 2456–2464.
[25] Y. Yuan, X. Zhang, L. Liu, X. Jiang, J. Lv, Z. Li, Z. Zou, Synthesis and photocatalytic characterization of a new photocatalyst $BaZrO_3$, Int. J. Hydrogen Energy 33 (2008) 5941–5949.
[26] Z. Khan, M. Qureshi, Tantalum doped BaZrO3 for efficient photocatalytic hydrogen generation by water splitting, Catal. Commun. 28 (2012) 82–85.
[27] Y. Mao, J. He, X. Sun, W. Li, X. Lu, J. Gana, Electrochemical synthesis of hierarchical Cu_2O stars with enhanced photoelectrochemical properties, Electrochim. Acta 62 (2012) 1–7.
[28] S. Parida, S.K. Rout, S. Cavalcante, E. Sinha, M. Siu, Structural refinement, optical and microwave dielectric properties of $BaZrO_3$, Ceram. Int. 38 (2012) 2129–2138.
[29] T. Kosugi, S. Kanek, Novel spray-pyrolysis deposition of cuprous oxide thin films, J. Am. Ceram. Soc. 81 (1998) 3117–3124.
[30] M.L. Moreira, M.F.C. Gurgel, G.P. Mambrini, P.S. Pizani, J.A. Varela, Photoluminescence of barium titanate and barium zirconate in multilayer disordered thin films at room temperature, J. Phys. Chem. A 112 (2008) 8938–8942.
[31] B. Alfonso, B. Stefano, R. Raffaele, Band offsets in lattice-matched heterojunctions: a model and first-principles calculations for GaAs/AlAs, Phys. Rev. Lett. 61 (1988) 734–738.
[32] G. Van de Walle Chris, R.M. Martin, Theoretical calculations of heterojunction discontinuities in the Si/Ge system, Phys. Rev. B 34 (1986) 5621–5635.
[33] L. Colombo, R. Resta, S. Baroni, Valence-band offsets at strained Si/Ge interfaces, Phys. Rev. B 44 (1991) 5572–5578.
[34] M. Peressi, N. Binggeli, A. Baldereschi, Band engineering at interfaces: theory and numerical experiments, Phys. D. Appl. Phys. 31 (1998) 1273–1278.
[35] R. Shaltaf, G.M. Rignanese, X. Gonze, F. Giustino, A. Pasquarello, Band offsets at the Si/SiO_2 interface from many-body perturbation theory, Phys. Rev. Lett. 100 (2008) 186401.
[36] M. French, R. Schwartz, H. Stolz, R. Redmer, Electronic band structure of Cu_2O by spin density functional theory, J. Phys. Condens. Matter 2 (1) (2009) 1–6.
[37] W.Y. Ching, N. Yong Xu, W.K. Wong, Ground-state and optical properties of Cu_2O and CuO crystals, Phys. Rev. B 40 (1989) 7684–7692.
[38] M. Zemzemi, S. Alaya, Band offset of the ZnO/Cu_2O heterojunction from ab initio calculations, Superlattices Microstruct. 64 (2001) 311–318.
[39] K. Anandan, V. Rajendran, Size controlled synthesis of SnO_2 nanoparticles: facile solvothermal process, J. Non-Oxide Glasses 2 (2010) 83–89.
[40] A. Tamirat, Rick John, A. Dubale, Su Wei-Nien, Hwang Bing-Joe, Using hematite for photoelectrochemical water splitting: a review of current progress and challenges, Nanoscale Horiz. 1 (2016) 243–267.
[41] D. Sharma, S. Upadhayay, V.R. Satsangi, R. Shrivastav, U.V. Waghmare, S. Dass, Improved photoelectrochemical water splitting performance of $Cu_2O/SrTiO_3$ heterojunction photoelectrode, J. Phys. Chem. C 118 (2014) 25320–25329.
[42] S. Poulston, P.M. Parlett, P. Stone, M. Bowker, surface oxidation and reduction of CuO and CuzO studied using XPS and XAES, Surf. Interface Anal. 24 (1996) 811–820.
[43] L. Huanga, F. Peng, S. Fumio, Ohuchi, In situ XPS study of band structures at Cu_2O/TiO_2 heterojunctions interface, Surf. Sci. 603 (2009) 2825–2834.
[44] H. Zhou, Y. Mao, S.S. Wong, Shape control and spectroscopy of crystalline $BaZrO_3$ perovskite particles, J. Mater. Chem. 17 (2007) 1707–1713.
[45] T.C. Dang, D.L. Pham, H.C. Le, V.H. Pham, TiO_2/CdS nanocomposite films: fabrication, characterization, electronic and optical properties, Adv. Nat. Sci. Nanosci. Nanotechnol. 1 (2010) 2043–6254.
[46] J.K. Kim, K. Shin, D. Hwan, Yoo P.J. Wang, G.Y. Han, J.K. Park, Hematite modified tungsten trioxide nanoparticle photoanode for solar water oxidation, J. Power Source 210 (2012) 32–37.
[47] D. Sharma, A. Verma, V.R. Satsangi, R. Shrivastav, U.V. Waghmare, S. Dass, Nanostructured $BaTiO_3/Cu_2O$ heterojunction with improved photoelectrochemical activity for H_2 evolution: experimental and first-principles analysis, Appl. Catal. B Environ. 189 (2016) 75–85.
[48] A. Verma, A. Srivastav, D. Sharma, A. Banerjee, S.V.R. Sharma, R. Shrivastav, D.K. Avasthi, S. Dass, A study on the effect of low energy ion beam irradiation on Au/TiO_2 system for its application in photoelectrochemical splitting of water, Nucl. Inst. Methods Phys. Res. B 379 (2016) 255–261.
[49] X. Lin, J. Xing, W. Wang, Z. Shan, F. Xu, F. Huang, Photocatalytic activities of heterojunction semiconductors $Bi_2O_3/BaTiO_3$: a strategy for the design of efficient combined photocatalysts, J. Phys. Chem. C 111 (2007) 18288–18293.
[50] W. Wang, J. Wang, Z. Wang, X. Wei, L. Liu, Q. Ren, W. Gao, Y. Lianga, H. Shi, p–n junction CuO/$BiVO_4$ heterogeneous nanostructures: synthesis and highly efficient visible-light photocatalytic performance, Dalton Trans. 43 (2014) 6735–6744.
[51] W. Yuan, J. Yuan, J. Xie, C.M. Li, Polymer-mediated self-assembly of TiO2@ Cu2O Core–Shell nanowire array for highly efficient photoelectrochemical water oxidation, ACS Appl. Mater. Interfaces 8 (2016) 6082–6092.
[52] Y.J. Chi, H.G. Fu, L.H. Qi, K.Y. Shi, H.B. Zhang, H.T. Yu, Preparation and photoelectric performance of ITO/TiO_2/CdS composite thin films, J. Photochem. Photobiol. A 195 (2008) 357–363.
[53] N.T. Hahn, H. Ye, D.W. Flaherty, A.J. Bard, C.B. Mullins, Reactive ballistic deposition of alpha-Fe2O3 thin films for photoelectrochemical water oxidation, ACS Nano 4 (2010) 1977–1986.
[54] D. Sharma, S. Upadhyay, A. Verma, V.R. Satsangi, R. Shrivastav, S. Dass, Nanostructured Ti-Fe_2O_3/Cu_2O heterojunction photoelectrode for efficient hydrogen production, Thin Solid Films 574 (2015) 125–133.

Wireless Pers Commun
DOI 10.1007/s11277-017-4370-5

Energy Efficient Probabilistic Clustering Technique for Data Aggregation in Wireless Sensor Network

Rajesh K. Yadav[1] · **Daya Gupta**[1] · **D. K. Lobiyal**[2]

© Springer Science+Business Media New York 2017

Abstract Economic utilization of energy in wireless sensor network, composed of tiny battery powered sensor nodes constrained in energy and computation power is a critical issue. Clustering techniques are most often used to reduce the consumption of energy by the sensor nodes due to data transmission. A widely used class of clustering techniques is probabilistic clustering in which a predetermined optimal probability is used to facilitate the cluster head selection process. This paper aims to devise a technique that improves the energy efficiency of probabilistic clustering algorithms by optimizing the number of clusters and the distribution of cluster heads in the network. We also present two generic approaches to integrate proposed technique into the existing probabilistic clustering algorithms. The simulation results show a considerable improvement in energy efficiency of probabilistic clustering protocols and consequently a prolonged network life time.

Keywords Clustering · Data aggregation · Probabilistic · Sensor network · Energy efficiency · Epoch

✉ Rajesh K. Yadav
rkyadav@dce.edu

Daya Gupta
dgupta@dce.ac.in

D. K. Lobiyal
lobiyal@gmail.com

[1] Department of Computer Engineering, Delhi Technological University, New Delhi, India

[2] School of Computer and System Sciences, Jawaharlal Nehru University, New Delhi, India

Published online: 27 May 2017

1 Introduction

The use of wireless sensor network has grown enormously in last decade. Wireless sensor networks are used in variety of applications such as monitoring physical and environmental conditions, military surveillance, live stock tracking, home appliance monitoring etc. In most wireless sensor network (WSN) applications nowadays the entire network must have the ability to operate unattended in harsh environments in which pure human access and monitoring cannot be easily scheduled or efficiently managed or it's even not feasible at all [1, 2]. There is a crucial need for scalable and energy efficient routing and data gathering and aggregation protocols in corresponding large-scale environments. Sensor nodes are generally deployed randomly uncontrolled means in many significant wireless sensor network and they form a network in an Ad hoc manner [3, 4]. Sensors in such networks are battery powered and energy constrained and their batteries usually cannot be recharged. Therefore we need energy aware routing and data gathering protocols that offer high scalability and low energy consumption for a long network lifetime. In wireless sensor network hundreds to thousands of sensor nodes are deployed usually randomly. Sensors sense their environment and send their sensed data to a processing centre, called as "Sink" or "Base Station" where all the data is collected and processed. Many routing algorithms have been proposed for efficient transmission of data between base station and sensor nodes. Grouping of sensor nodes into clusters has been widely used by researchers to satisfy the scalability, high energy efficiency and prolong network lifetime objectives [5–8]. In clustering the whole sensor network is partitioned into multiple groups of sensor nodes. Each group is called a cluster and each cluster has a leader called cluster head that perform special tasks such as data aggregation and fusion. Clustering has number of other benefits and corresponding objectives, In addition to supporting network scalability and decreasing energy consumption through data aggregation. The route setup can be localize within the cluster and thus reduce the size of the routing table stored at the individual sensor node. Clustering also conserve communication bandwidth by limiting the scope of inter-cluster communication among Cluster heads (CHs) and reduces redundant exchange of information among sensor nodes. Clustered architecture is use full for sensor networks because of its inherent suitability for data fusion the data gathered by all members of cluster can be fused at cluster head. A Cluster head can schedule activities in a cluster so that node can switch to low power sleep mode and reduce the energy consumption [1, 3, 8].

Paper organization: The rest of the paper is organized in following manner. Section 2 describes some representative protocols in the category of probabilistic clustering. Section 3 describes in detail our proposed modifications in existing clustering algorithms. Section 4 provides performance evaluation and simulation results. Section 5 concludes the whole work with a direction to future work.

2 Related Work

There is a widely used class of clustering algorithms called probabilistic clustering algorithms. The main aim of these algorithms was to prolong the network life time and reduce the energy consumption. There are number of clustering algorithms have been specifically designed for WSNs [9–14]. These proposed clustering techniques widely vary depending on the node deployment and bootstrapping schemes, the pursued network architecture, the characteristics of the cluster head (CH) nodes and the network operation model. Low

energy adaptive clustering hierarchy protocol improves using fuzzy logic, which takes battery level, distance, and node density into consideration [15, 16]. In Energy Efficient Strong Head clustering (EESH) [17], nodes are promoted CHs according to their respective residual energies, their respective degrees and the distance to and the residual energy of their neighbors. A CH may be elected by the sensors in a cluster or pre-assigned by the network designer. A CH may also be just one of the sensors or a node that is richer in resources. Most significant and widely used representatives of this category are LEACH [18], HEED [19], SEP [20], EEHC [21], TEEN [22].

2.1 Low Energy Adaptive Clustering Hierarchy (LEACH)

LEACH is one of the most popular clustering protocols proposed for WSNs. LEACH uses rotation of cluster heads to balance network energy consumption. The operation of the LEACH is divided into a number of rounds. Each round includes a set-up phase and a steady phase. Clusters are organised in set-up phase while sensed data are transferred from sensors to cluster heads in steady phase. In LEACH, cluster heads are elected in a distributed manner and independent from each other. Nodes make their decision to become a cluster heads by choosing a random number r in interval [0, 1]. A sensor node decides to be a CH in a round if its generated random number is less than the below a threshold [18]:

$$T(n) = \begin{cases} \dfrac{p}{1 - p\left(rmod\dfrac{1}{p}\right)} & if \quad n \in G \\ 0 & otherwise \end{cases} \tag{1}$$

where p is the desired percentage of CH nodes in the sensor population, r is the current round number; G is the set of nodes that have not been CHs in the last $1/p$ rounds.

2.2 Hybrid Energy Efficient Distributed Clustering (HEED)

HEED is another improved and very popular energy efficient protocol. It allows multi hop communication among cluster heads (CHs) and base station (BS). HEED is a hierarchical, distributed, clustering scheme in which a one-hop communication mechanism retained within each cluster [19]. There are two basic parameters, residual energy and intra cluster communication cost to choose CH nodes. The initial set of CHs obtained by using residual energy of nodes. On the other hand, intra-cluster communication cost reflects the node degree or node's proximity to the neighbour and is used by the nodes in deciding to join a cluster or not. Thus, unlike LEACH, it does not select Cluster head nodes randomly. Only sensors that have a high residual energy can become CH nodes.

2.3 Stable Election Protocol (SEP)

SEP is a heterogeneous proactive protocol proposed in [20] with two levels of heterogeneity. It considers two types of sensor nodes, normal nodes and advance nodes. Advance nodes have more energy as compared to normal nodes. In SEP both types of nodes have weighted probabilities of becoming cluster heads. Advance nodes have more chances of becoming cluster head than normal nodes. Weighted election probabilities of the normal and advance nodes, P_{nrm} and P_{adv} can be calculated as:

$$P_{nrm} = \frac{p}{1 + \alpha m} \tag{2}$$

$$P_{adv} = \frac{p(1 + \alpha)}{1 + \alpha m} \tag{3}$$

2.4 Energy Efficient Hierarchical Clustering (EEHC)

EEHC is another probabilistic clustering protocol proposed in [21]. LEACH protocol allowed only one level of clustering but EEHC allow more than one level of clustering. This protocol works in a bottom up fashion. It is a distributed algorithm for organization of sensor nodes k-hop hierarchical cluster. Initially, each sensor node is elected as a CH with probability '*p*' and announces its election to the neighbouring nodes within its communication range. The above CHs are now called the 'volunteer' CHs. Next, all the nodes that are within '*k*'-hops distance from a 'volunteer' CH, are supposed to receive the election message either directly or through intermediate forwarding. Consequently, any node that receives such Cluster head election message and is not itself a Cluster head (CH), becomes a member of the closest cluster. Additionally, a number of forced Cluster heads are elected from nodes that are neither CHs nor belong to a cluster. Specifically, if the election messages do not reach a node within a preset time interval *t*, the node becomes a 'forced' CH assuming that it is not within 'k' hops of all volunteer CHs.

2.5 Threshold Sensitive Energy Efficient Network Protocol (TEEN)

This protocol focuses on information aggregation rather than on cluster formation. It provides a hierarchical clustered structure, grouping nearby nodes within the same cluster. TEEN [22] use data centric method with hierarchical approach. The protocol defines two thresholds: the hard threshold is a threshold (absolute) value for the sensed attribute, while the soft threshold is a threshold (small change) value of the sensed attribute. The soft threshold can be varied, depending of the criticality of the sensed attribute and the target application. The nodes transmit sensor readings only when they fall above the hard threshold and change by given amount (soft threshold). In TEEN, medium continuously sensed by sensor nodes, but data transmission is done less frequently which favours the energy saving. However, if the thresholds are not crossed, the nodes will never communicate. TEEN does not support periodic reports.

3 Proposed Technique

In this section the detailed reasoning of proposed Energy Efficient Probabilistic Clustering Technique (EEPCT) for Data Aggregation in Wireless Sensor Network is discussed. We have used the same radio energy model for power consumption in radio transmission as used in most of the probabilistic Clustering algorithms such as LEACH [18].

3.1 Requirement of Clusters of Almost Same Size

Probabilistic clustering techniques usually do not take into account the relative positions of selected cluster heads; as a result there is a fair chance that in several rounds a considerable

number of the CHs will be either in proximity or very far from each other. In case, when a considerable number of cluster heads are in close proximity, the number of nodes in the clusters associated with these cluster heads will be considerably low in comparison to other clusters which will give rise to uneven clustering. While a well balanced clustering i.e. nearly same sized clusters is crucial for reasonable performance of probabilistic clustering techniques such as LEACH, TEEN etc. On the contrary if selected cluster heads are very far from each other, there will be a significant wastage of energy in inter-cluster communication in protocols such as HEED, EEHC etc.

Therefore a uniform distribution of cluster heads over entire network region is desirable to achieve a well balanced clustering.

So as our first proposed improvement, we want to introduce a parameter ξ to denote the closeness and its value depends on the size, node density and number of cluster heads to be selected. If in a particular round the distance between any two selected cluster heads is smaller than ξ, they will be considered too close to each other and one of them has to drop its decision of becoming cluster head (CH).

3.2 Maintaining the Optimal Percentage of CHs in Each Round

Though the above proposed improvement in cluster head selection process causes a significant improvement in the performance of probabilistic clustering algorithms, it has a small drawback. Consider the case when a large number of potential CHs (the nodes for which the value of generated random number is less than the threshold value) in a round of a particular epoch are in close proximity then many of them will drop their decision to become CH. In such a situation the number of selected cluster heads would be significantly lesser than the optimal number of cluster heads. This will lead to bigger size clusters and consequently more energy consumption in intra-cluster communication. Also in the last round of the epoch all the nodes that have not become cluster head so far in that particular epoch will have to become cluster head and so chances will be higher for the selection of more cluster heads than optimal number of cluster heads. This will increase the long distance transmission to the sink. To deal with this situation we want to increase the number of potential cluster heads in each round so that the number of selected CHs after dropping some of potential CHs because of ξ-closeness be as near to optimal value as possible. For this, we need to raise the threshold value in Eq. (1) so that more nodes than usual will be eligible to become cluster head.

So as our second proposed improvement another parameter σ introduced, denoting the threshold increment factor. The increment in the threshold value given by Eq. (1) will not be same for each round in an epoch as original threshold value itself increase in each subsequent round of an epoch and becomes equal to 1 in last round but we want to keep σ constant in each round so we can calculate new increased threshold using

$$T_{NEW}(n) = T(n) + \sigma(1 - T(n)) \tag{4}$$

In our assumption σ will be constant for a particular configuration of network and a particular choice of closeness factor ξ.

3.3 Choosing Closeness Factor ξ and Threshold Increment Factor σ

The values of ξ and σ are very crucial for better performance of clustering algorithms. Closeness Factor ξ depends on the network configuration, i.e. size of the network, number

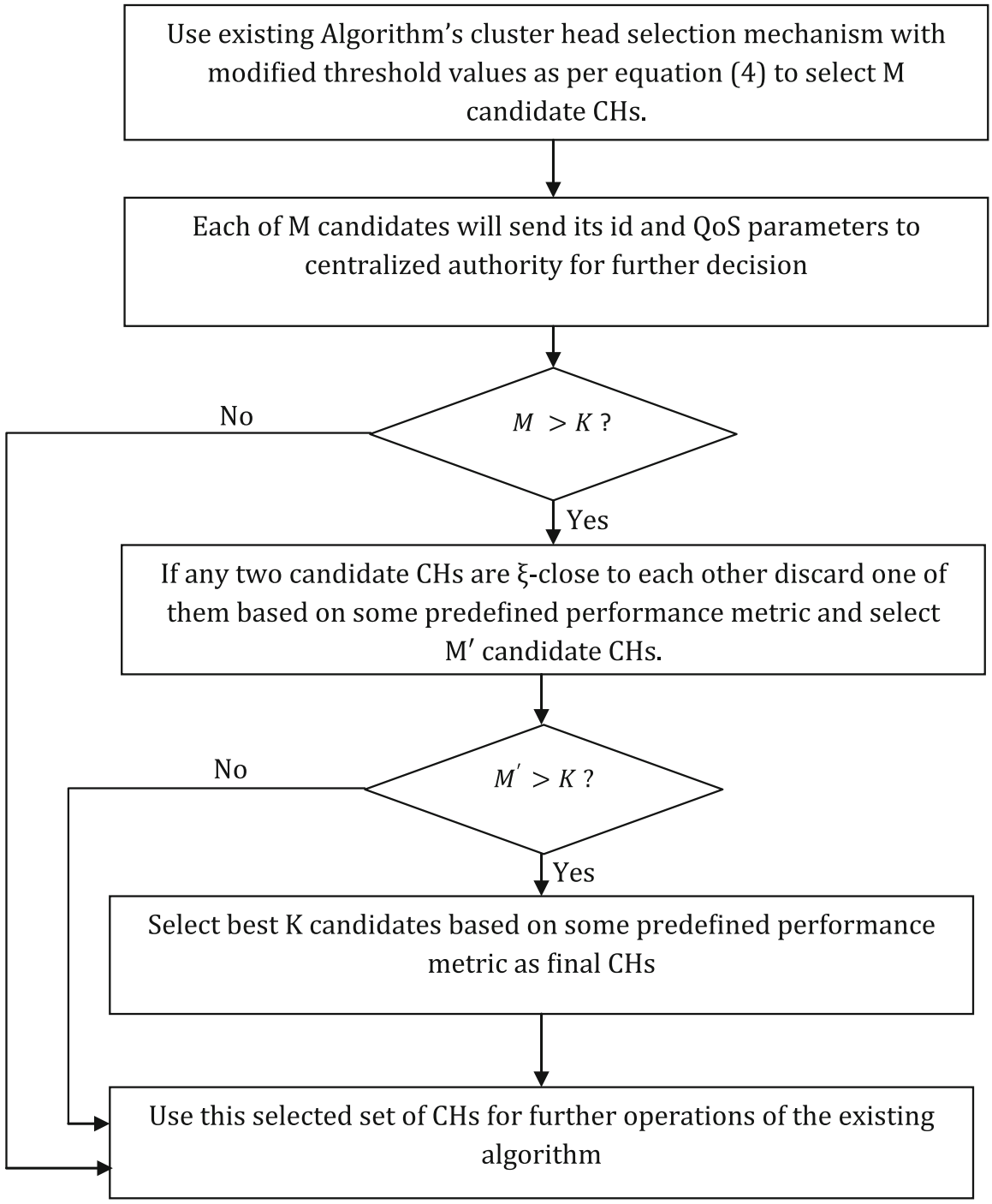


Fig. 1 A centralized approach for integrating proposed improvements to existing algorithms

of nodes deployed in the network, density of nodes, and optimal number of cluster heads, while the value of σ depends on the value of ξ as well as on network configuration. We have calculated the optimal values of ξ and σ experimentally by running multiple instances of the algorithm for different combinations of ξ and σ, keeping other parameters of the network constant.

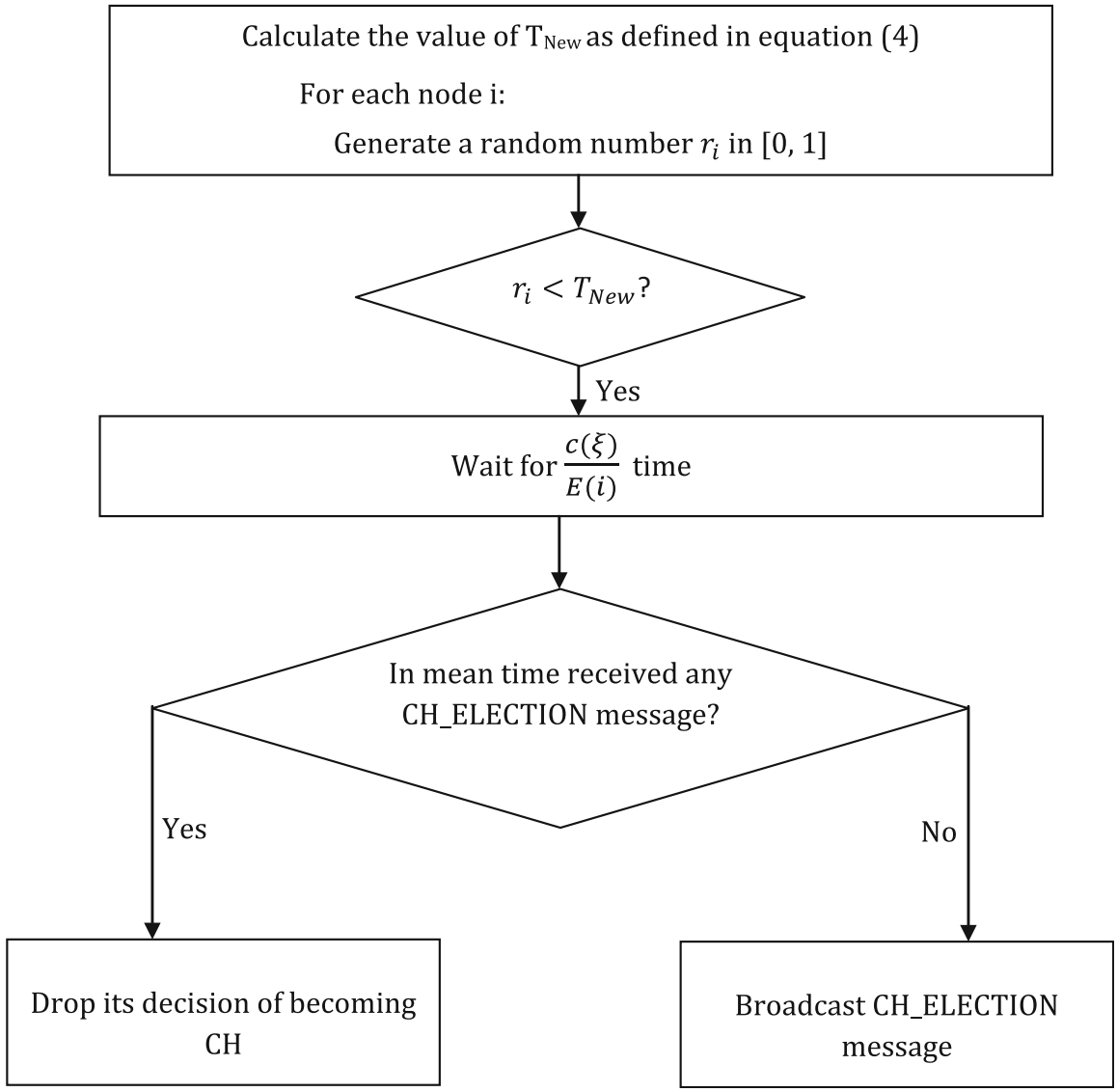


Fig. 2 A distributed approach for integrating proposed improvements to existing algorithms

Table 1 Simulation Parameters

Description	Parameter	Value
Initial energy	E_0	0.5 J
Electronic circuitry energy	E_{elec}	50 nJ/bit
Multi-path co-efficient	ε_{mp}	10 pJ/bit/m2
Free space co-efficient	ε_{fs}	0.0013 pJ/bit/m4
Data aggregation energy	E_{DA}	5 nJ/bit/signal
Data packet size	P_{pkt}	30 bytes
Total no. of nodes	N	100
Optimal percentage of CHs	P_{opt}	0.1

3.4 Integration with the Existing Probabilistic Techniques

The above two proposed improvements integrated with existing probabilistic clustering techniques to improve their energy efficiency. We have used two generic approaches of

Table 2 Determination of ξ and σ

ξ	σ	Energy consumption in configuration 1 (J)	Energy consumption in configuration 2 (J)
0	0.00	14.69	119.65
5	0.05	14.56	117.74
5	0.09	14.66	117.80
5	0.17	14.27	119.46
10	0.05	13.62	119.43
10	0.15	13.73	116.77
10	0.20	13.91	115.99
11	0.05	13.54	115.54
11	0.15	13.46	115.54
11	0.19	13.27	115.54
12	0.05	13.22	115.59
12	0.15	13.47	114.91
12	0.20	13.67	115.67
13	0.05	13.22	115.27
13	0.15	13.08	115.13
13	0.20	13.17	115.17
14	0.05	13.10	115.76
14	0.15	13.01	115. 58
14	0.20	12.93	115.58
15	0.00	13.23	116.44
15	0.05	13.01	116.01
15	0.10	12.66	116.62
15	0.11	12.61	116.49
15	0.13	12.63	116.49
15	0.15	12.54	116.32
15	0.16	12.51	117.51
15	0.17	12.84	117.96
15	0.19	12.73	116.59
16	0.05	12.87	116.87
16	0.15	12.69	116.76
16	0.20	12.67	116.14
17	0.05	12.95	116.97
17	0.15	12.83	116.52
17	0.25	12.83	117.01
18	0.05	12.87	116.66
18	0.10	12.67	117.67
18	0.20	12.68	116.19
19	0.05	12.93	114.93
19	0.15	13.11	114.27
19	0.20	13.09	114.81
20	0.05	13.83	115.36

Table 2 continued

ξ	σ	Energy consumption in configuration 1 (J)	Energy consumption in configuration 2 (J)
20	0.20	13.59	116.94
20	0.25	13.77	116.27
22	0.05	13.69	115.07
22	0.10	13.62	115.21
22	0.15	13.71	114.71
25	0.05	14.20	112.52
25	0.15	14.25	112.65
25	0.20	14.19	111.19
27	0.05	14.30	111.69
27	0.15	14.67	112.74
27	0.25	14.83	111.73
29	0.02	14.55	110.88
29	0.04	14.07	110.86
29	0.05	14.34	110.47
29	0.07	14.39	110.39
29	0.09	14.94	110.43
29	0.10	14.99	110.59
29	0.13	14.21	110.04
29	0.14	14.69	110.54
29	0.15	15.12	110.87
29	0.16	14.99	110.51
29	0.20	14.45	112.68
29	0.21	14.56	112.37
29	0.25	14.74	114.90
30	0.05	14.69	114.12
30	0.10	15.67	115.07
30	0.20	15.19	115.43

integration, one that requires little intervention of centralized authority such as base station and other follows a completely a distributed approach.

1. Centralized Approach: In Fig. 1 functioning of centralized approach of integration discussed.
   Here K is the optimal number of CHs that is precalculated using analytical methods as described in [18]. The predefined performance metric can include QoS factors such as quality of link between candidate and base station, congestion, node density around the candidate node, its residual energy etc.
2. Distributed Approach: In distributed approach each node autonomously takes decision of becoming a cluster head or not without intervention of any central authority. The following Fig. 2 explains a distributed approach for integration.
   Here E(i) is the residual energy of node i and c(ξ) is a function whose value depends only on closeness factor ξ. Waiting for $\frac{c(\xi)}{E(i)}$ time ensures that among all the candidate CHs that are in ξ-closeness range of each other only the node with highest residual

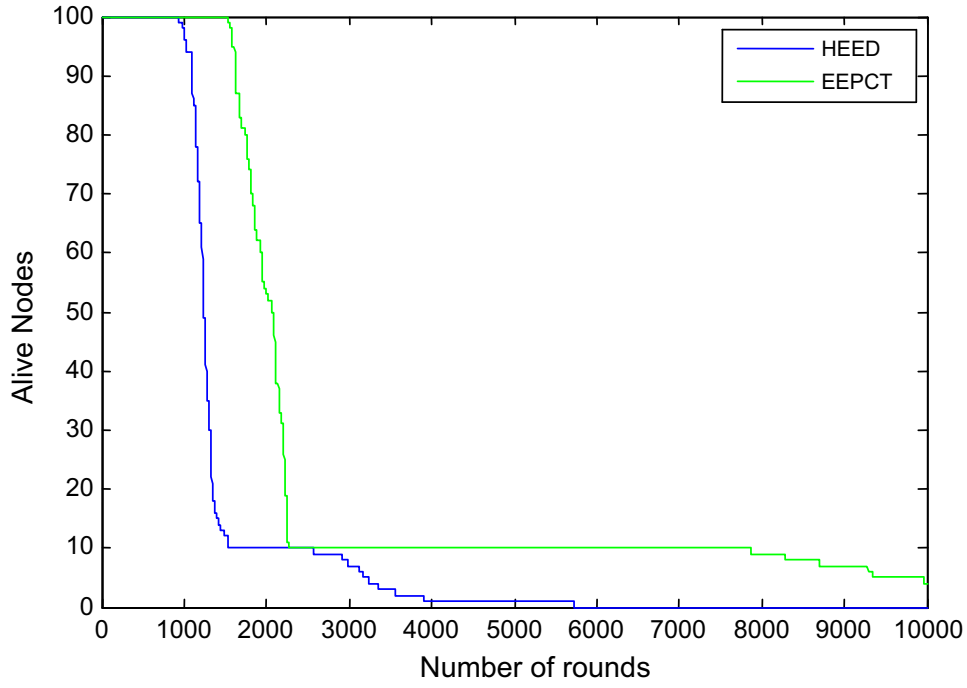


Fig. 3 Number of alive nodes per round

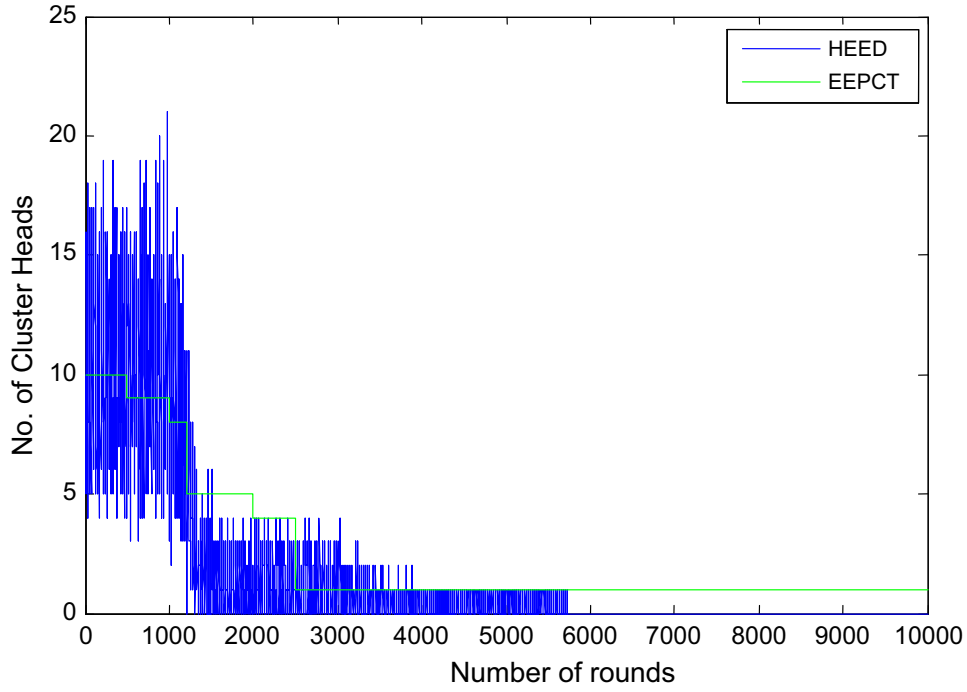


Fig. 4 Number of Cluster heads (CHs) per round

energy will be chosen as cluster head (CH).

Both centralized and distributed approaches have their own advantages. Centralized approach offers better QoS implementation mechanism and more tolerant to variance in values of σ and ξ from optimal value. On the other hand distributed approach requires minimal message exchange among nodes for CH election. Here we consider c as a function of ξ only but for QoS implementation we can make c as a function of other QoS parameters with ξ as well.

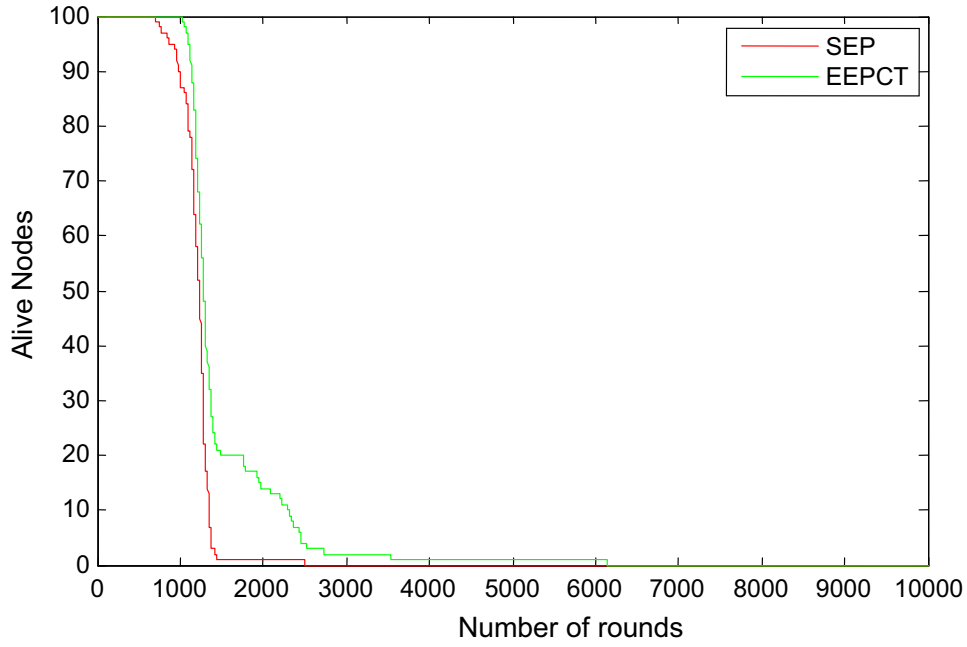


Fig. 5 Number of alive nodes per round

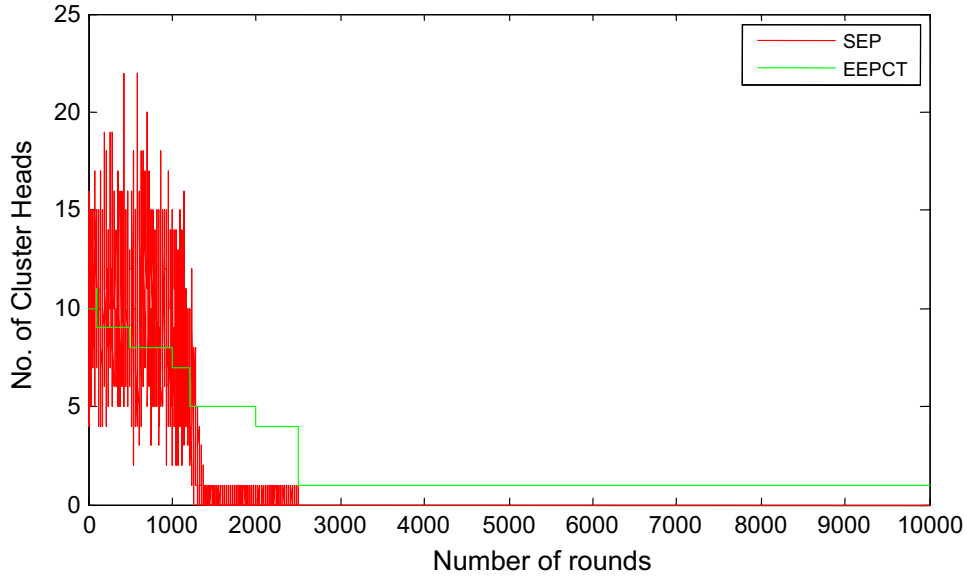


Fig. 6 Number of Cluster heads (CHs) per round

4 Simulation Results

We conducted a rigorous simulation study to evaluate the performance of our proposed Energy Efficient Probabilistic Clustering Technique (EEPCT). We assume a square network field of dimension 100 m $\times$ 100 m with 100 sensors deployed in it and MATLAB is used for the purpose of simulation. Specific parameters for simulation are shown in Table 1. The performance of EEPCT evaluated and compared with some representative probabilistic clustering algorithms such as LEACH [18], HEED [19], SEP [20] and TEEN [22].

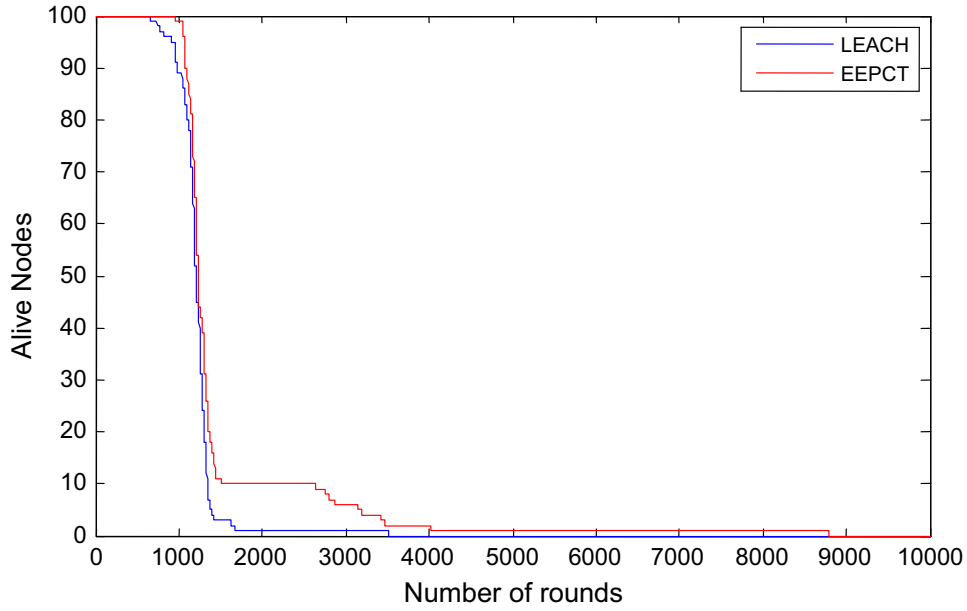


Fig. 7 Number of alive nodes per round

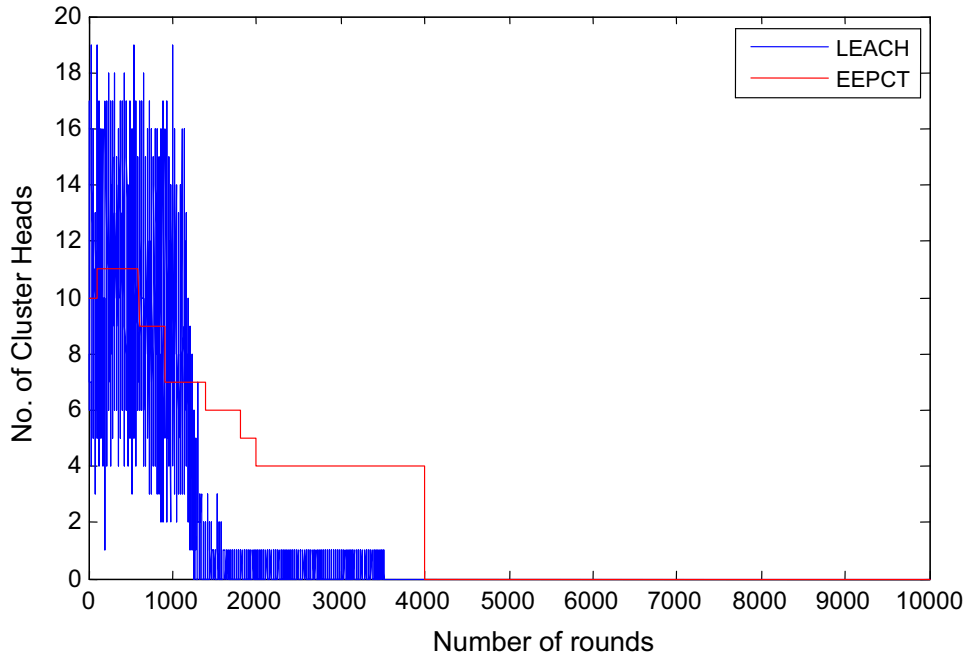


Fig. 8 Number of Cluster heads (CHs) per round

4.1 Determination of ξ and σ

To determine optimal values of ξ and σ we run our algorithm multiple times for different combinations of ξ and σ for a given configuration of network and choose best values. Table 2 shows energy consumption for different combinations of ξ and σ for two different network configurations.

We have calculated the optimal value of ξ and σ for two Network configurations. In first configuration 30 sensors deployed in a 50 m × 50 m network field. In second configuration we have taken a 100 m × 100 m network field with 100 sensors deployed in it. We

found 15 m and 0.16 to be optimal values for ξ and σ respectively for first configuration. Similarly for second configuration optimal values are 29 m and 0.13 respectively.

4.2 Performance Evaluation

In above figures we have shown the performance of existing probabilistic clustering algorithms in their original form and after the application of our proposed technique. We can see in Figs. 3 and 4 that after applying our proposed technique there is an increase of 43% in stability period of HEED and an overall 66% increment in total lifetime of network in simulated environment.

Similarly in case of SEP, in Figs. 5 and 6 we can see a considerable increase in the energy efficiency and network life time. We can also observe that after application of our proposed modification there is significant improvement in the no. of clusters formed in each round.

Similarly in LEACH, there is a significant improvement of in overall lifetime of the network. There is also an increase of around 37% in stability period of Network (Figs. 7, 8).

5 Conclusion

In this paper we proposed Energy Efficient Probabilistic Clustering Technique (EEPCT) with two changes in existing probabilistic clustering algorithms to ensure more or less optimal number of clusters and well distributed CHs across the network in each round. We also used two ways to apply proposed technique to existing clustering algorithms. Simulation results show that proposed technique obtains considerable improvement and allows a large stable network lifetime compared to LEACH, SEP, and HEED. In this paper we focused on probabilistic clustering algorithms but in future it can be extended to other class of clustering Algorithms.

References

1. Abbasi, A. A., & Younis, M. (2007). A survey on clustering algorithms for wireless sensor networks. *Computer Communications, 30,* 2826–2841.
2. Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirici, E. (2002). Wireless sensor network: A survey. *Computer Networks, 38*(4), 393–422.
3. Sohrabi, K., et al. (2000). Protocols for self-organization of a wireless sensor network. *IEEE Personal Communications, 7*(5), 16–27.
4. Min, R., et al. (2001). Low power wireless sensor networks. In *Proceedings of International Conference on VLSI Design*, Bangalore, India.
5. Dilip, K., Trilok, C. A., & Patel, R. B. (2009). EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks. *Computer Communications, 32*(4), 662–667.
6. Zhou, H., Wu, Y., Hu, Y., & Xie, G. (2010). A novel stable selection and reliable transmission protocol for clustered heterogeneous wireless sensor networks. *Computer Communications, 33*(15), 1843–1849.
7. Attea, B. A., & Khalil, E. A. (2012). A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks. *Applied Soft Computing Journal, 12*(7), 1950–1957.
8. Liliana, M., Arboleda, C., Nidal, N. (2006). Comparison of clustering algorithms and protocols for wireless sensor networks. In *Proceedings of IEEE CCECE/CCGEI Conference*, Ottawa, Ontario, Canada, pp. 1787–1792.

9. Gupta, G., Younis, M. (2003). Load-balanced clustering in wireless sensor networks. In: *Proceedings of the International Conference on Communication (ICC 2003)*, Anchorage, Alaska.
10. Bandyopadhyay, S., Coyle, E. (2003). An energy efficient hierarchical clustering algorithm for wireless sensor networks. In *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, San Francisco, California.
11. Ghiasi, S., Srivastava, A., Yang, X., & Sarrafzadeh, M. (2004). Optimal energy aware clustering in sensor networks. *Sensors Magazine MDPI, 1*(1), 258–269.
12. Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). Application specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Networking, 1,* 660–670.
13. Islam, A. A., Hyder, C. S., Kabir, H., & Naznin, M. (2010). Finding the optimal percentage of cluster heads from a new and complete mathematical model on leach. *Wireless Sensor Network, 2*(2), 129–140.
14. Wei, D., Kaplan, S., Chan, H. A. (2008). Energy efficient clustering algorithms for wireless sensor networks. In *Proceedings of IEEE Communications Society (ICC 2008)*, pp. 236–240.
15. Zhang, D., et al. (2014). An energy-balanced routing method based on forward-aware factor for wireless sensor network. *IEEE Transactions on Industrial Informatics, 10*(1), 766–773.
16. Ran, G., Zhang, H., & Gong, S. (2010). Improving on LEACH protocol of wireless sensor networks using fuzzy logic. *Journal of Information & Computational Science, 7,* 767–775.
17 Zhou, W., Chen, H. M., & Zhang, X. F. (2007). An energy efficient strong head clustering algorithm for wireless sensor networks. In *2007 international conference on wireless communications, networking and mobile computing, WiCOM 2007*, pp. 2584–2587.
18. Heinzelman, W. R., Chandrakasan, A. P., & Balakrishnan, H. (2000). Energy efficient communication protocol for wireless micro sensor networks. In *Proceedings of the 33rd Hawaaian Interantional Conference on System Sciences.*
19. Younis, O., & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing, 3*(4), 366–379.
20. Smaragdakis, G., Matta, I. & Bestavros, A. (2004). SEP: A stable election protocol for clustered heterogeneous wireless sensor networks. In *Proceedings of the International Workshop on SANPA*, pp. 251–261.
21. Bandyopadhyay, S. & Coyle, E. (2003). An energy efficient hierarchical clustering algorithm for wireless sensor networks. In *22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, San Francisco, CA.
22. Manjeshwar, A. & Agrawal, D. P. (2001). TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In *Proceedings of the 15th International Parallel and Distributed Processing Symposium*, San Francisco, CA.



Rajesh K. Yadav presently working as an Assistant Professor in the Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India. His areas of research interest are Mobile ad-hoc Networks and Wireless Sensor Networks.



Daya Gupta received her Ph.D. from IIT, Delhi. Presently working as a Professor in the Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India. Her areas of research interest are Image Processing and Artificial intelligence.



D. K. Lobiyal received his Ph.D. from Jawaharlal Nehru University, New Delhi, India. Presently, he is working as a Professor in the School of Computer and Systems Sciences at Jawaharlal Nehru University. His areas of research interest are Mobile ad-hoc Networks, Wireless Sensor Networks and Natural Language Processing (NLP).

Face Detection and Extraction from Low Resolution Surveillance Video Using Motion Segmentation

Vikram Mutneja[1]
Ph.D. Research Scholar,
I.K. Gujral Punjab Technical University, Kapurthala, Punjab (India)
vikram.mutneja@gmail.com

Dr. Satvir Singh[2],
Associate Professor,
I.K. Gujral Punjab Technical University Main Campus, Kapurthala, Punjab (India)
drsatvir.in@gmail.com

***Abstract*—Face detection is a prominent research domain in the field of digital image processing particularly in the field of video surveillance systems. Today is the world of video technology starting from low resolution videos to the high definition videos. The videos obtained from surveillance systems are often of low resolution due to the reasons such as distance between the camera and place of footage, environment factors, wide coverage area, installation problems, out of focus, bandwidth issue, hardware constraints, storage space limitations etc. because of which the frames need to be compressed or converted to lower resolution before storage. In this paper, we have worked on motion segmentation based face detection from low resolution surveillance videos. The motion segmentation is used to extract the region of interest from the current frame. Thereafter only the pixels obtained after the motion segmentation are subjected to the face detection process. The haar features based face detection has been used in this work, employing the image scaling to facilitate multi-scale face detection. Considerable search space reduction and efficiency boost has been achieved by proposed motion segmentation technique.**

***Keywords*-**Face Detection, Low Resolution Surveillance Videos, Motion Segmentation, Haar Features, AdaBoost

I. INTRODUCTION

Video surveillance systems also known as CCTV (Closed circuit television) are the systems which use video camera for the purpose of surveillance. These systems keep on recording the video footage of the scene continuously 24 hours a day and keep storage of video as per the storage capacity of the system e.g. of last one week or one month. Video surveillance is in use since quite a long time for the monitoring and security purposes in various public and private places such as railway stations, offices, banks, roads, showrooms and shopping malls. The method of monitoring had been manual most of the times which usually is prone to human errors on account of factors such as fatigue, lack of human attention. The advent of technology and computational power has given way to transformation of traditional manual surveillance systems into intelligent video surveillance systems, which not only record the data, but also do the intelligent i.e. automatic monitoring of the video. Intelligent surveillance video systems aim at two main functions. Firstly detection of the objects of interest e.g. people, vehicles. Secondly tracking, activity analysis and recognition of the objects for the events detection and recognition e.g. detection and tracking of human face and head for the analysis of human attention, detection of a "person smoking" activity which involves detection of human, hands, cigarette and smoking activity.

There is great thrust in research in the field of automatic surveillance video systems. The proposed work is concerning detection and extraction of human faces from the low resolution surveillance video sequence. In the surveillance video systems where task is to detect, track and recognize people as well as analyze people activities, detection and extraction of human faces is of paramount importance. It is very important to attach the identity to persons being detected and tracked in the video. From the fact that human faces are used as biometric entity, human faces are generally used to attach identity to a detected human in the surveillance video. Detecting the human faces in surveillance videos is a challenging task on account of various factors such as illumination, low resolution of surveillance cameras, human pose, face pose, facial gestures, head pose, face occlusions such as goggles, scarf, face hair and various face and head accessories. The facial images in surveillance videos are of very low resolution, therefore it is challenging task to detect the faces from low resolution surveillance videos.

The training and detection framework used in this work has been derived from our previous work related to modified haar features and AdaBoost based face detection system Mutneja and Singh (2017). This paper is structured as follows: Section 2 provides Literature Survey, Section 3 gives details of proposed technique for the face detection algorithm, Section 4 explains experimental setup, Section 5 discusses experimental results, and finally Section 6 discusses conclusions and future Scope.

II. LITERATURE SURVEY

Sarkar et al. (2012) worked on multiple face detection and tracking from low resolution video sequence, they used skin color information for face region estimation, eyes and mouth region localization in the detected skin region to confirm it as face region. Chen et al. (2007) used video object and skin color segmentation for face localization and neural networks for face quality analysis. Kasturi et al. (2009) proposed very robust framework for performance evaluation of face detection and tracking in surveillance videos. Zhu and Ramanan (2012) presented a unified model for face detection, pose estimation and landmarks localization in real world cluttered images. Wang (2014) proposed a complete algorithmic description, a learning code and a learned face detector that can be applied to any color image. Since the Viola-Jones algorithm typically gives multiple detections, a post-processing step is also proposed to reduce detection redundancy using a robustness argument.

Zakaria and Suandi (2011) used combination of neural network and adaboost, Huang et al. (2011) used combination of Genetic algorithm and neural network, Martinez-Gonzalez and Ayala-Ramirez (2011) used neural networks for real time face detection, Jaisakthi and Aravindan (2011), used data and sensor fusion technique using SVM, Guan et al. (2012) proposed face localization using fuzzy classifier, haar features and YcbCr color features, Pan et al. (2013) used combination of haar like, local binary patterns and speeded up robust features is conjunction with SVM and PSO for multi-view face detection, Hiremath et al. (2012) implemented fuzzy geometric face model for searching face region using prominent face features such as eyes and mouth to detect faces. Seyedarabi et al. (2009) used skin color and face edge information to develop a fuzzy rule based classifier to extract head candidate from image using YcbCr colour space model. Ming Ouhyoung et al. (2012)used real time depth sensors for nose detection for human face localization to overcome face occulsions. Kuo et al. (2010) used fuzzy c-means for color recognition of objects in surveillance videos.

During the past two decades of research in the field of face detection, Viola and Jones (2004) did seminal work in face detection, key contributions of their work were: new image representation called integral image to facilitate faster calculation of features, adaboost learning algorithm for classifiers, cascade classifiers for faster computation. As per surveyed by Belaroussi and Milgram (2012), growing research field is concentrating in developing appearance based models for multi-view and rotation invariant face detections. In case of color image sequence, using skin color results in faster face localization and poses estimation.

Alionte and Lazar (2015) proposes practical implementation of a face detector based on Viola-Jones algorithm using Matlab cascade object detector is presented. Employing the system type object vision.CascadeObjectDetector, eight face detectors were developed using the trainCascadeObjectDetector function and tuning the number of cascade layer and the False Alarm Rate. For different tuning parameters, the performances of the face detectors were analyzed.

III. PROPOSED ALGORITHM

The face detection process in proposed work is composed of multiple steps. Firstly the difference of current video frame is found with respect to the previous frame to find the motion segmented pixels. Further only the motion segmented pixels are subjected to the generation of sub-images, to be parsed for the face classification. The size of the sub-images is same as the detector window. The detector has been trained using the example images of size 18×18. The configuration of the trained detector being used is as in table 1.

TABLE 1: DETECTOR CONFIGURATION

Training Images Size	18×18
Haar Features Pool Size	32384
Number of Weak Classifiers:	1147
Number of Stages in Cascade	19
Configuration	[2, 2, 3, 5, 5, 10, 20, 30, 40, 50, 60, 80, 90, 100, 110, 120, 130, 140, 150]

The multi-scale face detection has been facilitated by applying image scaling. The minimum and maximum values of scaling factors are calculated based upon the size of input image, minimum size of face to be detected and maximum image size which can be handled by system. Algorithm 1 shows the working of the face detection process and Algorithm 2 shows the function to process the sub-images generated from the motion segmented pixels called from the previous algorithm.

Algorithm 1 Face Detection Process

1: Load the First Frame of the Video ($pFrame$), Size of the Video Frame, $NR \times NC$
2: Compute the
$LowestScalingFactor(LR)$,
$UpperScalingfactor(UR)$,
$N = NumberofScalingFactors$
3: Load the Haar Classifiers of the Detection Cascade Kernel
4: Initialize and load the function (ProcessMotion) for performing the detection on motion segmented pixels
5: Loop1: Load the Current Frame of Video ($cFrame$)
6: Extract the motion segmented pixels as vectors MX and MY, which are x,y co-ordinates respectively of motion segmented region
7: Initialize the Array to Hold the Locations of Detected Faces
8: Loop2: For multi-scale face detection, i=1:N
9: Scale Image and Motion Pixel Vectors XM, YM: $img1 = Img * i * LR, XMi = XM * i * LR, YMi = YM * i * LR$
10: Determine $xMax$, $yMax$ Maximum pixel locations for sub-image generation based upon size of scaled image, $xMin = yMin = 2$.
11: Eliminate the the values from XMi and YMi which are out of range $xMintoxMax$ and $yMintoyMax$
12: Initialize the array to hold the sub-images to be generated.
13: Loop3: To scan for pixel locations from XMi and YMi for sub-images generation with shift interval $SL(i)$
14: if X, Y lies in the previous detected face area **then**
15: Skip Rest and Go to Next Motion Pixel Location (Loop3)
16: Create the sub-image, compute its integral image and add it to array of sub-images (Integral Versions)
17: Iterate Loop2 for Next Motion Pixel Location
18: Call the function (Pass the argument: Generated array of sub-images (Integral Versions) to detect faces from them)
19: Update Array of Detected Faces for Scaling Factor=$i * LR$, Iterate Loop2 for Next Scaling Factor
20: All Faces Detected in Current Frame,
21: Set $pFrame = cFrame$, i.e. Previous frame=Current frame and Iterate Loop1 for Next Frame

IV. Experimental Setup

The machine running windows 8.1 (64 Bit) on Intel core i3 1.9 GHz has been used to test the proposed method. The work has been done in MATLAB version 8.2.0.701 (R2013b).

V. Results and Discussion

The proposed algorithm has been tested on the test videos from the low resolution surveillance systems from the dataset INRIA (2004) with specifications: Frame Width: 384 pixels, Frame Height: 288 pixels, Frame Rate: 25 Frames per Second, Data Rate: 1184 kbps, Total Bitrate: 1184 kbps. Figure 1 and Figure 2 shows the result of motion segmentation on few of the frames of test videos from dataset INRIA.

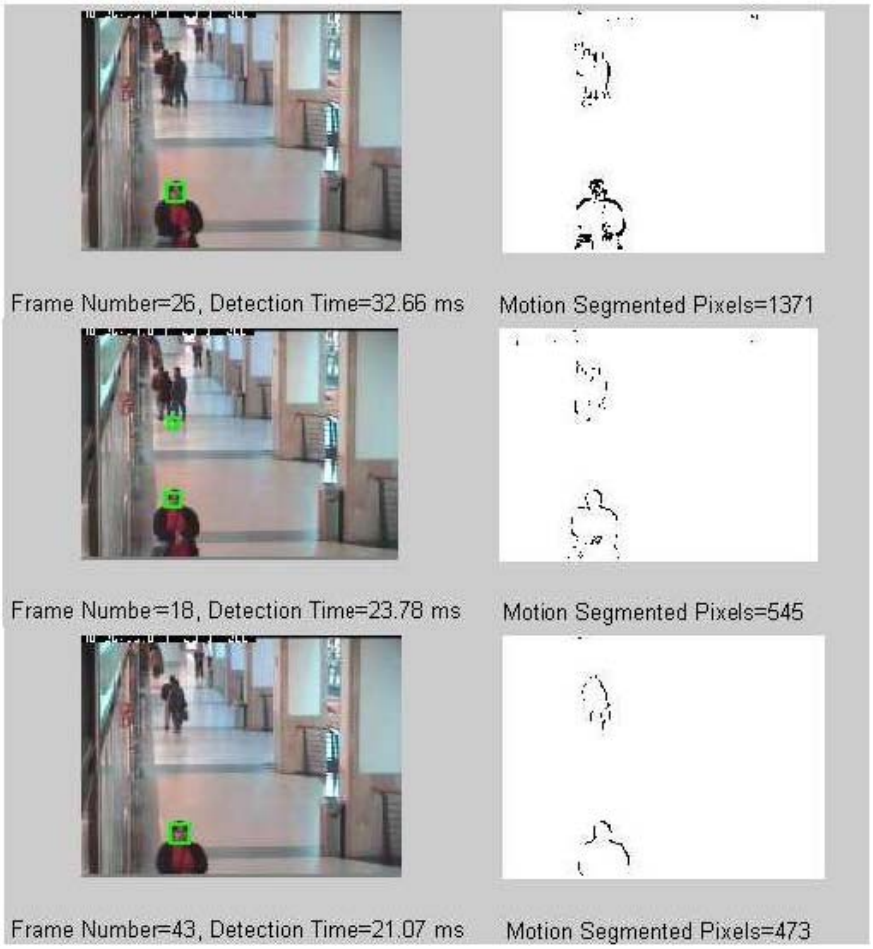


Figure 1: Result Motion Segmentation on a Test Video (OneLeaveShop1cor.mpg) From Dataset INRIA

TABLE 2: PROCESSING TIME AND NUMBER OF SEGMENTED PIXELS OF STARTING 20 FRAMES (TEST VIDEO 1)

Frame	Detection Time (ms)	Segmented Pixels
1	0.00	0
2	85.24	315
3	22.65	252
4	20.68	368
5	22.59	529
6	22.27	632
7	27.79	736
8	23.64	663
9	21.74	546
10	18.47	531
11	19.50	485
12	21.16	568
13	21.76	577
14	23.20	677
15	0.00	0
16	33.50	1387
17	24.81	725
18	26.51	893
19	24.30	669
20	22.53	552

Table 2 shows the time cost and number of pixels segmented by proposed motion segmentation method on a test video1 ("OneLeaveShopReenter2cor.mpg" from INRIA). We have been able to achieve the speed of processing of the order of 37.86 fps, with the average detection rate of 98.86% at the false acceptance rate 4.44 %.



Figure 2: Detection Results on Few Test Video Frames

Table 3 shows the time cost and number of pixels segmented on a test video 2("OneShopOneWait1cor.mpg" from INRIA) by proposed motion segmentation method. We have been able to achieve the speed of processing on the test video 2 of the order of 54.57 fps, with the average detection rate of 99.52% at the false acceptance rate 4.21 %. From the results, it is inferred that we have been able to achieve the promising results in terms of processing speed as well as detection accuracy on the test videos. Further by comparison of results of first and second video, it is inferred that better time efficiency has been achieved in the second video because of less number of motion segmented pixels to be processed.

Algorithm 2 Function to Process Sub-Images for Face Detection Process (ProcessMotion)

1: Load Variables/Function Arguments $DetectorWindowSize(DWS)$, $NumberofWeakClassifiers(NCF)$, $NumberofStagesofCascadeKernel(numLevels)$, $ArrayofFeatures(F)$, $Thresholds(TH)$ and $NumberofSub-Images(numSubImages)$
2: Loop1: for Iterating the Sub-images, For sIdx=1:$numSubImages$
3: Compute Absolute Base Address for pixel Operations: $BaseAddress = sIdx * DWS * DWS$
4: Loop2: for Feature Number ($fNum$), for fNum=1:NCF
5: Initialize array to hold binary results of decision stumps, $face_or_noface(NCF)$
6: Fetching the co-ordinates of sub-rectangles from feature array and compute Haar Feature Value
7: Compare it for Lower and Upper Bounds of Thresholds
8: Update $face_or_noface$ Array
9: Iterate Loop2 for evaluating all the Features
10: Loop3: For i=1:$numLevels$; Operation of Detection Cascade
11: Threshold value $V(i)$ computed of Stage number i
12: **if** $if(V(i) > TH(i))$ **then**
13: Go to Loop3, to Iterate to Next Stage of Detection Cascade
14: **else**
15: $status(sIdx) = 0$, Detected as Non_Face **return**
16: **if** $i == numLevels$ **then** $status(sIdx) = 1$ (Detected as Face) **return**
17: Iterate Loop1 for Next Sub-image

TABLE 3: PROCESSING TIME AND NUMBER OF SEGMENTED PIXELS OF STARTING 20 FRAMES (TEST VIDEO 2)

Frame	Detection Time (ms)	Segmented Pixels
1	0.00	0
2	81.51	25
3	18.37	52
4	14.04	51
5	13.64	21
6	13.82	40
7	13.92	19
8	14.25	41
9	14.33	33
10	14.97	19
11	14.46	31
12	14.58	17
13	15.35	28
14	13.60	24
15	14.67	10
16	13.36	27
17	13.56	15
18	15.19	53
19	14.58	28
20	14.38	8

VI. CONCLUSION AND FUTURE SCOPE

In the proposed system, we have developed the algorithm for the detection and extraction of faces from low resolution surveillance videos using motion segmentation based technique. The inter-frame difference is performed and the pixels having the difference above the set threshold level are marked as motion segmented pixels, which are subjected to generation of sub-images for the detection of faces. The detector of small size i.e. 18×18 has been used so as to target the low resolution faces. The testing has been performed on the test videos from the dataset INRIA. From the results achieved, we contend that proposed algorithm is very effective in detection of faces from low resolution surveillance videos.

We intend to do further improvements in the proposed algorithm by incorporating the handling of faces with severe head orientations and occlusions. We further want to integrate the proposed algorithm with the complete video surveillance based face biometric system and acceleration with the help of GPU computing.

ACKNOWLEDGMENT

The proposed work has been carried under the research related to first author's Ph.D. in field of facial image processing from low resolution surveillance videos, registered as part time research scholar under I.K. Gujral Punjab Technical University, Kapurthala, Punjab (India).

REFERENCES

[1] Elena Alionte and Corneliu Lazar. A practical implementation of face detection by using matlab cascade object detector. In System Theory, Control and Computing (ICSTCC), 2015 19th International Conference on, pages 785–790. IEEE, 2015.
[2] Rachid Belaroussi and Maurice Milgram. A comparative study on face detection and tracking algorithms. Expert systems with Applications, 39 (8): 7158–7164, 2012.
[3] Tse-Wei Chen, Shou-Chieh Hsu, and Shao-Yi Chien. Automatic feature-based face scoring in surveillance systems. In Multimedia, 2007. ISM 2007. Ninth IEEE International Symposium on, pages 139–146. IEEE, 2007.
[4] Chen-Ning Guan, Chia-Feng Juang, and Guo-Cyuan Chen. Face localization using fuzzy classifier with wavelet-localized focus color features and shape features. Digital Signal Processing, 22 (6): 961–970, 2012.
[5] PS Hiremath, Manjunath Hiremath, and R Mahesh. Face detection and tracking in video sequence using fuzzy geometric face model and motion estimation. International Journal of Computer Applications, 58 (15), 2012.
[6] Chen-rong Huang, Jia-li Tang, and Yi-jun Liu. A new face detection method with ga-bp neural network. In Wireless Communications, Networking and Mobile Computing (WiCOM), 2011 7th International Conference on, pages 1–4. IEEE, 2011.
[7] EC Funded CAVIAR project/IST 2001 37540 INRIA. 2004. URL http://homepages.inf.ed.ac.uk/rbf/CAVIAR/.
[8] SM Jaisakthi and Chandrabose Aravindan. Face detection using data and sensor fusion techniques. In Soft Computing and Pattern Recognition (SoCPaR), 2011 International Conference of, pages 274–279. IEEE, 2011.
[9] Rangachar Kasturi, Dmitry Goldgof, Padmanabhan Soundararajan, Vasant Manohar, John Garofolo, Rachel Bowers, Matthew Boonstra, Valentina Korzhova, and Jing Zhang. Framework for performance evaluation of face, text, and vehicle detection and tracking in video: Data, metrics, and protocol. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31 (2): 319–336, 2009.
[10] Jong Yih Kuo, Tai Yu Lai, Fu-Chu Huang, and Kevin Liu. The color recognition of objects of survey and implementation on real-time video surveillance. In Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on, pages 3741–3748. IEEE, 2010.
[11] Angel Noe Martinez-Gonzalez and Victor Ayala-Ramirez. Real time face detection using neural networks. In Artificial Intelligence (MICAI), 2011 10th Mexican International Conference on, pages 144–149. IEEE, 2011.
[12] Ming Ouhyoung, Hong-Shiang Lin, Yi-Ting Wu, Yi-Shan Cheng, and Dominik Seifert. Unconventional approaches for facial animation and tracking. In SIGGRAPH Asia 2012 Technical Briefs, page 24. ACM, 2012.
[13] Hong Pan, Yaping Zhu, and Liangzheng Xia. Efficient and accurate face detection using heterogeneous feature descriptors and feature selection. Computer Vision and Image Understanding, 117 (1): 12–28, 2013.
[14] Rajib Sarkar, Sambit Bakshi, and Pankaj K Sa. A real-time model for multiple human face tracking from low-resolution surveillance videos. Procedia Technology, 6: 1004–1010, 2012.
[15] Hadi Seyedarabi, Saeed Mahdizadeh Bakhshmand, and Sohrab Khanmohammadi. Multi-pose head tracking using colour and edge features fuzzy aggregation for driver assistant system. In Signal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on, pages 385–390. IEEE, 2009.
[16] Paul Viola and Michael J Jones. Robust real-time face detection. International journal of computer vision, 57 (2): 137–154, 2004.
[17] Yi-Qing Wang. An analysis of the viola-jones face detection algorithm. Image Processing On Line, 4: 128–148, 2014.
[18] Zulhadi Zakaria and Shahrel A Suandi. Face detection using combination of neural network and adaboost. In TENCON 2011-2011 IEEE Region 10 Conference, pages 335–338. IEEE, 2011.
[19] Xiangxin Zhu and Deva Ramanan. Face detection, pose estimation, and landmark localization in the wild. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pages 2879–2886. IEEE, 2012.
[20] Vikram Mutneja and Satvir Singh. Modified viola–jones algorithm with gpu accelerated training and parallelized skin color filtering-based face detection. Journal of Real-Time Image Processing, pages 1–21, 2017.

AUTHORS PROFILE



Mr. Vikram Mutneja is Ph.D. research scholar in I.K. Gujral Punjab Technical University, Kapurthala (Punajb, India) in the discipline Electronics Engineering. He received his Bachelor's degree (B.Tech.) from Guru Nanak Dev University, Amritsar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Master's degree (M.Tech.) from Giani Zail Singh College of Engineering & Technology, Bathinda (Punjab) (India) under Punjab Technical University (PTU, Jalandhar) with first division in Electronics & Communication Engineering in year 2008. He is a Life Member of ISTE. He

has industrial experience of around 5 years (1998-2003) in which he worked in diverse fields such as Hardware and Networking Technical Support, Software Development for Web and Embedded Systems. During his teaching experience of around 13 years (2003-2016) he worked mainly in the areas of embedded systems, VLSI and Digital Signal Processing. He is working presently as Assistant Professor in department of Electronics & Communication Engineering in Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab (India). His current field of research includes areas of image processing for detection and extraction of facial features from low resolution surveillance videos, and acceleration using GPU computing. (Vikram.mutneja@gmail.com)



Dr. Satvir Singh is Associate Professor in I.K. Gujral Punjab Technical University (Main Campus), Kapurthala (Punajb, India). He received his Bachelor's degree (B.Tech.) from Dr. B. R. Ambedkar National Institute of Technology, Jalandhar, Punjab (India) with specialization in Electronics & Communication Engineering in year 1998, Master's degree (M.E.) from Delhi Technological University (Formerly, Delhi College of Engineering), Delhi (India) with distinction in Electronics & Communication Engineering in year 2000 and Doctoral degree (Ph.D.) from Maharshi Dayanand University, Rohtak, Haryana (India) in year 2011. He is a Life Member of ISTE and Corporate Member of IETE. During his 15 years of teaching experience he served as Assistant Professor and Head, Department of Electronics & Communication Engineering at BRCM College of Engineering & Technology, Bahal, (Bhiwani) Haryana, India and as Professor & Head, Department of Electronics & Communication Engineering at Shaheed Bhagat Singh State Technical Campus (Formerly, SBS College of Engineering & Technology), Ferozepur Punjab, India. His fields of special interest include Evolutionary Algorithms, High Performance Computing, Type-1 & Type-2 Fuzzy Logic Systems, Wireless Sensor Networks and Artificial Neural Networks for solving engineering problems. (Drsatvir.in@gmail.com)

emeraldinsight

Grey Systems: Theory and Application

Grey relational analysis coupled with principal component analysis for optimization of the cyclic parameters of solar driven organic rankine cycle

Deepak Tiwari, Ahmad Faizan Sherwani, MOHAMMAD ASJAD, Akhilesh Arora,

Article information:

To cite this document:
Deepak Tiwari, Ahmad Faizan Sherwani, MOHAMMAD ASJAD, Akhilesh Arora, "Grey relational analysis coupled with principal component analysis for optimization of the cyclic parameters of solar driven organic rankine cycle", Grey Systems: Theory and Application, https://doi.org/10.1108/GS-03-2017-0006
Permanent link to this document:
https://doi.org/10.1108/GS-03-2017-0006

Downloaded on: 02 July 2017, At: 21:22 (PT)
References: this document contains references to 0 other documents.
To copy this document: permissions@emeraldinsight.com
The fulltext of this document has been downloaded 3 times since 2017*
Access to this document was granted through an Emerald subscription provided by emerald-srm:544911 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

Grey relational analysis coupled with principal component analysis for optimization of the cyclic parameters of solar driven organic rankine cycle

Abstract:

Purpose: In this research work, the effect of four controllable parameters (Fuel mixture, Evaporation bubble point temperature, expander inlet temperature and condensation dew point temperature) of a Solar driven organic rankine cycle on First law efficiency, exergetic efficiency, exergy destruction and volume flow ratio (expander outlet /expander inlet) has been investigated.

Design/Methodology/Approch: Nine experiments as per Taguchi's standard L9 orthogonal array were performed on Solar driven organic rankine cycle. Subsequently, multi-response optimization was performed using grey relational and principal component analyses.

Findings: The results revealed that grey relational analysis along with the principal component analysis is a simple as well as effective method for solving the multi-response optimization problem and it provides the optimal combination of Solar driven organic rankine cycle parameters. Further, the analysis of variance (ANOVA) was also employed to identify the most significant parameter based on percentage of contribution of each cyclic parameter. Confirmation tests were performed to check the validity of the results which revealed good agreement between predicted and experimental values of the response variables at optimum combination of the input parameters.The optimal combination of process parameters is the set with A3 (the best fuel mixture in context to optimal performance is 0.9% Butane and 0.1% Pentane by weight), B2(Evaporation bubble point temperature=358K), C1(Condensation dew point temperature=300K), D3(Expander inlet temperature=370K).

Research limitations/implications: In this research work, the Taguchi based grey relational analysis coupled with PCA has been successfully carried out whereas, for any optimized solution, it is required to have a real time scenario that may be taken into consideration by the application of different soft computing techniques like genetic algorithm, Simulated annealing, etc. The result generated are purely based on theoretical modeling and for further research, experimental analyses are required to consolidate the generated results.

Originality/value: This piece of research work will be helpful to users of solar energy, academician, researchers, and other concerned persons, in understanding the importance, severity and benefits obtained by the application, implementation and optimization of the cyclic parameters of solar driven organic rankine cycle.

Keywords: Solar driven organic Rankine cycle, Taguchi method, Grey relational analysis, Principal component analysis, ANOVA

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Nomenclature

Abbreviation	
$\dot{E}D$	Exergy destruction rate
GWP	Global Warming Potential
ORC	Organic Rankine cycle
ODP	Ozone depletion Potential
PCA	Principal Component analysis

Symbols	
$\dot{m}$	Mass flow rate (kg/sec)
Q	Heat transfer (W)
S	Specific entropy (kj/kgK)
W	Work transfer (W)
η	Efficiency
Subscript	
cond	Condenser
e	Exit
h	Specific enthalpy(kj/kg)
H	Heat source
i	inlet
ev	Evaporator
ex	Exergetic
exp	Expander
H	Heat source
o	Ambient
P	Pump
R	Rejection
S	Supplied
T	Temperature
th	Thermal

1. Introduction

Nowadays, the perspective of accelerating energy demand, environmental pollutions, irregularities in climate change and depleting fossil fuels; the exploitation of energy from low grade renewable and sustainable energy heat sources (i.e.60-200^{o}C (Yamatto et al., 2001)) such as solar, geothermal, biomass and waste heat(i.e. byproduct of combustion) is overarching and inevitable. Among these, the solar energy is clean, abundant ($1.7x10^{4}$, about ten thousand times of present world's total energy consumption (Chen, 2010)), noiseless, cost effective among various other energy sources. The conversion of power from low grade heat source is difficult to achieve using conventional vapor powered Rankine cycle but the

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

power from low grade heat source is successfully converted using Organic Rankine cycle (Zhai, et al., 2016).

The working principle and structure of Organic Rankine cycle (ORC) is similar to conventional vapor power Rankine cycle but instead of water organic substance is used as working fluid in ORC, capable to operate at low temperature in power as well as combined power and heat application (Peris et al., 2015 and Kumar et al., 2016).

The lower thermal efficiency (DiPippo, 2004), lower exergetic efficiency (Sadeghi, et al., 2016), higher exergy destruction (Yamada, et al., 2011) in evaporator due to thermal mismatch between heat source and working fluid and appropriate working fluid selection (Chen, 2010) are limitations of ORC. Partial evaporating organic Rankine cycle (PEORC) (Zhou et al., 2016), Buoyancy organic Rankine cycle (Schoenmaker et al., 2011), Pumpless organic Rankine cycle (Yamada et al., 2013) and Organic flash cycle (Lee et al., 2016) are the recent advancements in field of an organic Rankine cycle.

Generally, the optimization techniques can be broadly classified into two i.e. traditional and non traditional. The traditional one is that uses the mathematics tools and techniques (Asjad, et al., 2013) while the non traditional may includes the artificial intelligence (for e.g., genetic algorithm, etc.(Khan et al, 2015)).

Taguchi technique is one of the traditional techniques that may use for parameter optimization in design of experiments. Keeping in view the inherent features of the said method to produce optimum results by performing lesser number of experiments and thereby consuming relatively lesser time and cost etc., it has been employed in the present study to determine the optimum parameters of a solar driven organic rankine cycle. Further, analysis of variance (ANOVA) has been used to determine significance of the input parameters. The conformity test is also conducted to check the robustness of the obtained results.

The paper is organized with section 2 covering the reviews of available literature, while section 3 develops and presents the methodology adopted to carry out the research. Section 4 presents the system description while section 5 develops the simulation process carried out to obtain the results. Section 6 presents the results whereas section 7 concludes the research presented in this work.

2. Literature review

There are some notable studies have been carried out and discussed based on low grade heat source powered organic Rankine cycle using pure fluids and their mixtures in recent years are as follows:

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Aghahosseini and Dincer (2013) performed thermodynamic analysis of low grade heat source powered organic Rankine cycle using pure and zeotropic mixtures and reported R600a, R245fa, R134a, R404a, R123 and R407c are potential working substances. Feng et al., (2015) proposed the thermo-economic multi objective optimization by Non-dominated sorting genetic algorithm and TOPSIS by considering two objective function maximum exergetic efficiency and minimum levelized energy cost for low grade waste heat source powered superheated ORC using 245fa/R227ea for various location of evaporation bubble and condensation dew point temperature.

Kumar et al., (2016) carried out parametric optimization for low grade waste heat powered using L9 orthogonal array of Taguchi's method and concluded that maximum thermal efficiency is obtained by using R-134a as working fluid at 393K expander inlet temperature,293K condenser temperature with 6kg/s mass flow rate. Lu et al., (2016) analyzed the organic Rankine cycle for set cooling water temperature increase and condensation bubble point temperature using R245fa/R600a and R601a/R600 zeotropic mixtures and concluded that temperature glide in mixtures enhances the net power output of the cycle.

Wang et al., (2013) carried out parametric optimization using genetic algorithm for solar powered recuperated organic Rankine cycle with flat plate collector and concluded that R123 and R245fa are better fluid than the R134a and isobutene. Wu et al., (2016) performed thermal and economic analysis of organic Rankine cycle using pure fluids and their mixtures of Butane / R245fa, RC318 / R245fa, R227ea / R245fa and concluded that the thermal performance of mixtures is better and inferior economical performance than pure fluid at certain mass fraction ratio.

Sadeghi et al., (2016) performed thermodynamic analysis and multi-purpose optimization by Pareto optimal frontier for simple ORC, Parallel two-stage ORC and series two stage ORC using R404A, R438A, R22M, R402A,R403B,R407A, R410A, R422D, R422A as zeotropic mixtures and reported that 27.76%, 24.98%, 24.79% enhancement in power generation and lower turbine size parameters as compared to pure fluid for simple ORC, series and parallel two stage organic Rankine cycle respectively.

Herberle et al., (2012) performed life time estimation for low grade geothermal heat source powered Organic Rankine cycle using isopentane,isobutane,propane,R123yf,R1233zd of low global warming potential are alternative working fluids to R245fa and R134a.

Herberle et al., (2016) carried out thermodynamic analysis for small-enthalpy (Temperature below 120°C) geothermal powered heat resource and concluded better performance of the

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

zeotropic working fluid than pure fluid. Andreasen et al., (2014) performed optimization of pure and mixtures working substance for low grade heat source powered using MATLAB genetic algorithm function and concluded that mixture of ethane/propane obtain 12.9% increment in net power.

Imran et al., (2016) conducted a parametric optimization of modified organic Rankine cycle by multipurpose MATLAB genetic algorithm to maximize exergetic efficiency and minimize specific investment cost using R245fa as a working fluid and concluded that the first law and exergetic efficiency is higher and higher investment cost than the basic ORC of regenerative and recuperative ORC respectively. Xu et al., (2015) evaluated the performance of direct steam generation of supercritical organic Rankine cycle using (LFR) solar concentrator and concluded that the net increase in efficiency is 19.53 of supercritical organic Rankine cycle using cyclohexane as working fluid.

Tchanche et al., (2009) performed thermodynamic analysis of solar powered organic Rankine cycle for twenty working fluid and reported that the first law efficiency lies between 3.7-4.9% and biggest part of exergy destruction occurred in evaporator (50%) and expander(13%), respectively.

From above aforementioned literature review it can be revealed that limited studies have been carried out for optimization of low grade heat source powered organic Rankine using zeotropic mixtures of low GWP and zero ODP. This paper describe the analysis methods and the optimization of the first law efficiency, exergetic efficiency and minimize the exergy destruction and volume flow ratio by the Grey analysis coupled with principal component analysis.

3. System description and Methodology

In this section, a description of the system used for analysis is presented along with the methodology. The next sub-sections present the system description.

3.1. System description

Solar energy is used as heat source in this study.Fig.1 exhibits the heat is supplied (2-3) to the working fluid in evaporator (collector) from global solar radiation which raises the temperature of the zeotropic mixture to the saturated vapor. The vapor enters into the expander (Turbine) where enthalpy drop takes place (3-4) is subsequently converted into the power. The heat from released superheated working fluid from expander is carried away through the cooling medium in condenser (5-1) which transform it into the saturated liquid which further pressurized by the pump(1-2) and thus completes the cycle.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

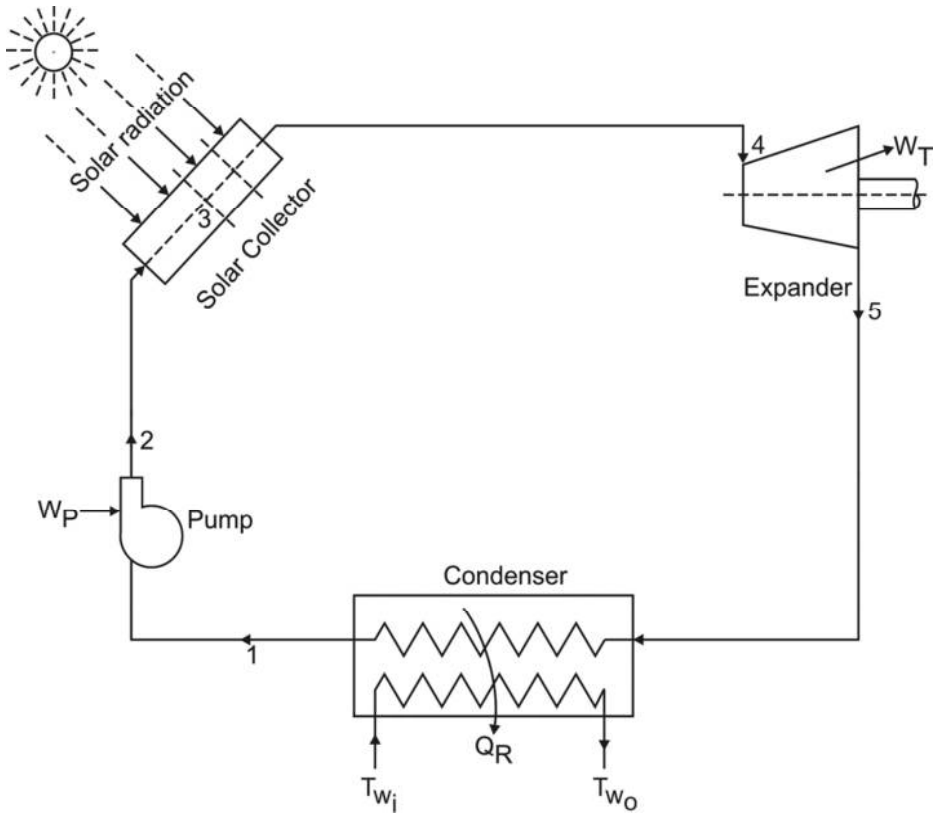


Figure 1 Basic organic Rankine cycle using solar heat source

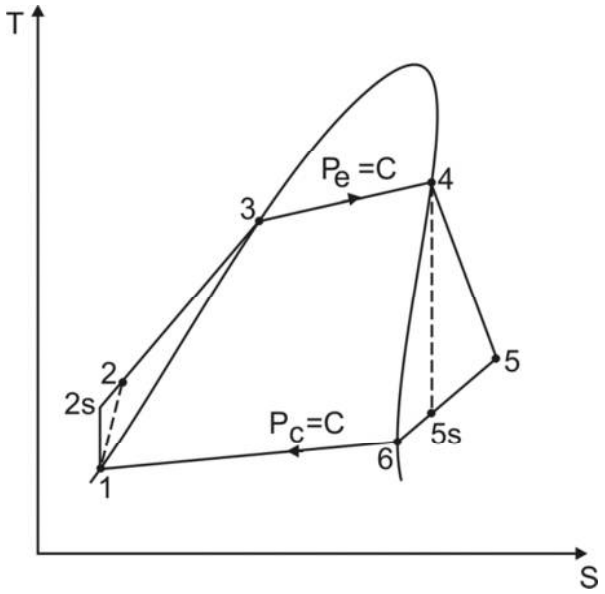


Fig.2.T-S diagram for organic Rankine cycle

Fig.2 Renders the T-s diagram for saturated Organic Rankine cycle using zeotropic mixtures. The proposed cycle described by following processes are as follows:

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Process2-3:Non isothermal constant pressure heat addition,Process:3-4s:isentropic expansion ,Process:5-1: non isothermal constant pressure heat rejection,Process:1-2s:isentropic compression.

The next subsection illustrates the methodology implemented to carry out the research.

3.2 Methodology

This section is further divided into three sub sections viz Taguchi's method, Grey Relational analysis and Principal component analysis respectively.

3.2.1 Taguchi's method

Taguchi is one of the powerful techniques used to design the best quality of the system when design parameters are discrete and quality wise. By this process, the optimization of characteristics of performance is done by assigning the designed parameters and alleviates the sensitivity of system's performance to the causes of variation. In Taguchi method system's parameters optimization is done by in three steps, parameter design, system design and tolerance design. Tolerance design is applicable to manufacturing field only it is not applicable in thermal field. Only parameter design is used in this study to optimize the characteristics of performance. Taguchi's method is widely applicable in manufacturing industries but till now limited works have been carried out for optimization of low grade heat source powered organic Rankine cycle. In this work parametric optimization is carried out by Taguchi's method for low grade solar powered ORC using low global warming potential and zero ODP zeotropic mixtures.

3.2.2 Grey Relation

Gray relation analysis is performed in following steps

1-Tabulate the results of experiment on the basis of plan of experiment.

2-The normalization of the data can be done in three ways viz; Higher the better (HB), Lower the better(LB) and Nominal the best(NB)(Gopalsamy et al., 2009; Selvaraj and Chandramohan, 2010).

The normalization to this sequence of original response

$$y_q(r) = \frac{max y_q(r) - y_q(r)}{\max y_q(r) - min y_q(r)} \quad (1)$$

$y_q(r)$ is the data obtained by normalization, for r^{th} response in q^{th} experiment.

3-The co-efficient of grey relation is determined by following relation

$$\varepsilon_q(\mathrm{r}) = \frac{\delta min + \omega \delta_{max}}{\delta_{or}(k) + \omega \delta_{max}} \quad (2)$$

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

ε_q(r) is coefficient of Grey Relationδ_{max} is highest value of $\delta_{or}(k)$, δmin is lowest value of $\delta_{or}(r)$ is absolute difference of $y_o(q)$ and y_q(q), $y_o(q)$ is ideal sequence.

$$\delta_{or}(r) = |y_o(q)\text{-}y_q(\mathrm{q})| \quad (3)$$

4-The grade of grey relation ($\wp_i$) is given by following equation

$$\wp_i = \frac{1}{\mathrm{m}}\sum_{\mathrm{q}=1}^{\mathrm{k}} \varepsilon_{\mathrm{q}}(\mathrm{r}) \quad (4)$$

Where k= is total responses and n is total output of responses.

3.2.3 PCA

Principal component analysis is type of statistical tool is used to remove extensity of data arrangement consist of plenty interrelating variables, although maintaining approximately same variation in data arrangement. It is discovered by Pearson (1901) and Chen(2010), although the foremost latest source is Jolliffe (2002).

Principal components analysis (PCA) is used to alter the data linearly to fresh coordinate arrangement are such as fresh arrangement of variables; principle components are dependent on real variables (Uncorrelated).The data of largest variance obtained by any data projection lies in first coordinate and the data of second largest variance in second coordinate and all.

If y is M dimension data taken from some static distribution condensed to 0 , F{x}= 0. The principal i^{th} component $y^Tb(i)$ of y is described through settled eigen's vector b(i) of covariance data matrix $B=F\{YY^T\}$ consist of greatest i^{th} value of eigen ʎ(i).The space within the space is linked with principal eigen's vector b(1),…….,b(N) (N<M) is known as principal component analysis of space within the space. Principal components analysis network is of neural apprehension of principal component analysis where neuran weight vector w(i) .

The weight vector matrix W=[w(1),w(2)…….w(m)] concentrate to principal eigen's vector b(i).

Principal component analysis is prominent used to calculate the optimal solution in various field. It is performed in following steps are given below

Step1-Maximize the linear altered variance $F\{[w(i)^Tx]^2\}$ or outcome of linear arrangement for orthonormal constrained $[w^Tw=I]$.

Step2- Minimize the mean of square rendition error $F[|y\text{-}ý|^2]$, $ý=ww^Ty$.

Step3-Unrelatedness of outcomes $w(i)^Tx$ for various neuron following orthonormalization transform$[ww^T=I]$.

Step4-Minimize the rendition of entropy.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

4. Thermodynamics Modeling of ORC system

The thermodynamic modeling of the ORC system based on principles of conservation of mass, balance of energy and exergy. The program code is written in MATLAB15a interfaced with REFPROP 9.0. The properties of zeotropic mixtures are taken from REFPROP 9.0. For simplification of cycle there are following assumptions are made.

1- All the processes in this cycle are under steady state

2- Neglect the heat loses and pressure drops in heat exchangers(Evaporator, Condenser) (Aghahosseini, S and Dincer, 2013)

3- Isentropic efficiency of turbine (η_t)=0.8 (Shu et al., 2014 and Sadeghi et al., 2016)

4-Isentropic efficiency of pump$\left(\eta_p\right) = 0.7$ (Kang et al., 2015; Sadeghi et al., 2016)

5-Evaporation bubble points (353,358,363K) (Deethayat et al., 2015)

6-Expander inlet temperature (360,365,370K) (Tchanche et al., 2009 and Andreasen et al., 2014)

6-Condensation dew point temperature (300,305,310K) [Deethayat et al.,2016 and Kang et al., 2015)

7-Ambient temperature (T_o) =20°C (Feng et al., 2015)

8-Mass flow rate of the working fluid (kg/s) =1 (Chen et al., 2011)

9-Heat source inlet temperature(T_h) =100°C (Abadi et al., 2015)

10-Working fluid considered is butane/pentane

11-Mass fractions of considered working fluid (0.1, 0.5, 0.9)

4.1 Energy and exergy balance

The energy balance equation gives heat and work transfer for ORC system components under steady state and neglecting steady state condition is as follows for open system

$q - w = \sum_e m_e h_e - \sum_i m h_i$ (Sadeghi et al., 2016) (5)

q=heat transfer rate

m=mass flow rate

q, w heat and work transfer rate respectively.

h_e ,h_i specific enthalpy at exit to inlet of the system respectively.

Expander (3-4)

The expander power output

$W_{exp} = \eta_{ex}(h_4 - h_{5s})$ (6)

Where W_{ex} is expander power ouput, $\eta_{ex}$is isentropic exapnder efficiency, h_4 and h_{5s} are enthalpy of expander inlet and outlet respectively.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Condenser (5-1)

Heat rejected by condenser

$$Q_R = h_5 - h_1 \quad (7)$$

Where Q_R heat rejected in condenser, h_5 and h_1 are the enthalpies of condenser inlet and outlet respectively.

Pump (1-2)

$$W_P = (h_{2s} - h_1) / \eta_P \quad (8)$$

Where $W_P$is pump work, $\eta_P$the isentropic pump efficiency

Heat supplied to evaporator

$$Q_s = h_4 - h_2 \quad (9)$$

Where Q_s is heat supplied to the working fluid in evaporator , h_4 and h_2 enthalpy expander inlet and outlet respectively.

The energy or thermal efficiency of system

$$\eta_{th} = (W_t - W_P) / Q_s \quad (10)$$

The exergy destruction rate is given by

$$\dot{ED}=\dot{m}[(h_{in}-h_o)-T_O(s_{in} - s_o)] \quad (11)$$

Where $\dot{m}$ is mass flow rate of working fluid, h_{in} and h_o enthalpy inlet to outlet respectively, T_o ambient temperature.

Exergy destruction for Expander

$$\dot{ED}_{exp}=\dot{m}[T_O(s_5 - s_4)] \quad (12)$$

Exergy destruction in condenser

$$\dot{ED}_{cond}=\dot{m}[T_O((S_1 - S_5) - (h_1 - h_5) / T_o] \quad (13)$$

Exergy destruction for pump

$$\dot{ED}_p=\dot{m}[T_o(S_2 - S_1)] \quad (14)$$

Exergy destruction in evaporator

$$\dot{ED}_{ev} = T_o[(S_4 - S_2) - (h_4 - h_2) / T_H] \quad (15)$$

The total exergy destruction is the sum of exergy destruction occurred in all ORC system's components.

$$\dot{ED}_{tot} = (\dot{ED}_{exp} + \dot{ED}_p + \dot{ED}_{cond} + \dot{ED}_{ev}) \quad (16)$$

Exergetic efficiency

$$\eta_{ex} = \frac{W_{exp}}{Q_S\left(1-\frac{T_O}{T_H}\right)+W_P} \quad (17)$$

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

4.2 Working fluid selection

In this research work, the butane/pentane mixture is used as a working fluid for organic rankine cycle to optimize the first law efficiency, exergetic efficiency and minimize the exergy destruction and volume flow ratio by the Grey relational analysis coupled with principal component analysis.

The selection of appropriate working fluid is not an easy task for low grade heat source powered ORC. Due to severe environmental pollutions, growing energy demand it is mandatory to use environmental friendly and thermodynamic efficient working fluid. Zeotropic mixtures exhibits isobaric non isothermal phase change provide better thermal match between heat source and sink which results in lower exergy destruction and better cycle performance than pure fluids (Zhao and Bao, 2014). But due to separation in evaporation process, unascertained thermo-physical properties and smaller co-efficient of convective heat transfer (Lemmon et al., 2002) mixtures have not been widely applied. There is no such a ideal working substance for all perspective such as first law efficiency, area of heat exchanger, mass and volumetric flow rate (Andreasen et al., 2014).In this study butane/pentane (R600/R601) zeotropic mixture of hydrocarbon is chosen as working fluid which has zero ODP and relatively low GWP and thermodynamic efficient (Guo et al., 2015) in comparisons to other working fluids are reported in literature review. But the flammability of hydrocarbon is bottleneck for effective utilization in ORCs applications.

Table1. Thermophysical properties of the working fluid (Guo et al., 2015 and Geng et al., 2016)

Working Fluid	Critical Temperature(K)	Critical Pressure(MPa)	Type of Fluid	Flammability	ODP	GWP
Butane	425.13	3.76	Dry	Flammable	0	~20
Pentane	469.7	3.37	Dry	Flammable	0	~20

4.3 Cycle parameters

On the basis of extensive literature survey the performance of ORC using zeotropic mixtures depend upon selection of zeotropic mixtures, mass fraction of zeotropic mixtures, boiling point temperature difference between mixture' components of zeotropic mixtures, expander inlet temperature, value and location of bubble and dew point temperature in condensation

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

and evaporation process respectively ,however among various parameters: Mass fraction of working fluid, Evaporation bubble point Temperature, Condensation dew point Temperature, Expander inlet Temperature are are chosen to optimize the ORC performance in this study.

(a)-Mass fraction of working fluid: Condensation and evaporation temperature glide, Energy and exergy efficiencies, power output and exergy destruction strongly depends upon mass fraction. At certain mass fraction energy and exergy efficiency, power output is maximum while the exergy destruction is minimum (Feng et al., 2015; Kumar et al., 2016; Wang et al., 2013 and Zhai et al., 2016).

(b)-Evaporation bubble point temperature: It is the temperature at which liquid working fluid first starts to boil [Source. Du Pont^{TM} Suva$^{@}$]. With increasing evaporation temperature power output, thermal efficiency, exergetic efficiency increases and exergy destruction decreases (Shu et al., 2014).

(c) Condensation dew point temperature: It is the temperature at which vapor working fluid first starts to condense [Source. Du Pont^{TM} Suva$^{@}$]. Higher condensation dew point temperature exhibits larger temperature glide leads to lower enthalpy drop in expander as a result lower output is developed (Li et al., 2014).

(d)Expander inlet temperature: With increasing expander inlet temperature the power output and energy efficiency increases because of larger enthalpy drop across the expander (Shu et al., 2014).

The next section presents the description about the results generated by simulation.

5 Simulation

Likewise equations given in section 4.3 a computer code is developed in MATLAB 2015a interfaced with REFPROP 9.0 (Lemmon et al., 2002).The required properties of selected zeotropic working fluids are such as; normal boiling point, critical temperature and pressure, triple point and density etc. are drawn from REFPROP. The state of all input parameters; temperature, pressure , enthalpy, entropy are defined with the help of quality (dryness fraction) and pressure of working fluid at different locations; sub-cooled region, saturated liquid, two phase (Liquid+ Vapor), saturated, and superheated region. The MATLAB written code is versatile can also be applicable for any two of arbitrary chosen pure working fluids to form zeotropic mixture. The separation phenomena and undefined thermo-physical properties of working fluids during phase change are the limitation of simulation study. The value of input parameters which affects ORC cycle performance using with zeotropic mixtures are; evaporation temperature glide, condensation dew point temperature, expander

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

inlet temperature and mass fraction. The performance of organic Rankine cycle are evaluated in term of fist law efficiency, exergetic efficiency, exergy destruction and volume flow ratio (expander outlet /expander inlet) at evaporation bubble point temperature (353, 358, 363K), condensation dew point temperature (300, 305, 310K), expander inlet temperature (360, 365, 370) and mass fraction (0.1/0.9, 0.5/0.5, 0.9/0.1) by selecting desired zeotropic mixture of pure working fluids. The state of the input property are defined in MATLAB interfaced with REFPROP;

Evaluated property=calling of thermo-physical properties('Properties required','specification1', value1, 'specification2', value2 working fluid 1, working fluid2, [Mass fraction of first working fluid Mass fraction of second working fluid])

6. Results, analysis and discussions

The following steps were followed to estimate the optimal combinations of solar driven process parameters based on grey relational analysis coupled with principal component analysis:

1. Identify the process parameters
2. Selection of Orthogonal Array - Taguchi Method
3. Parametric estimation of the output based on the selected design
4. S/N ratios for the experimental data were estimated
5. Normalized the S/N ratio
6. Estimation of grey relation coefficient
7. Application of PCA to estimate the weight and thereby the grey relational grades
8. Apply ANOVA
9. Conformity test, to check the robustness of the optimal parameters.

1. Identify the process parameters: Refer section 4 and Table 2, given below.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited

This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table 2: Process parameters and their levels

Cycle parameters	Symbols	Units	Levels		
			Level 1	Level 2	Level 3
Fuel Mixture (Butane + Pentane)	A	Ratio	0.1	0.5	0.9
Evaporation bubble point temperature	B	K (Degree Kelvin)	353	358	363
Condensation dew point temperature	C	K (Degree Kelvin)	300	305	310
Expander inlet temperature	D	K (Degree Kelvin)	360	365	370

2. Selection of Orthogonal Array - Taguchi Method

The second step in this method is to select an appropriate orthogonal array (OA) which is based on the number of input parameters and their levels involved in the investigation. Depending on the number of parameters and their levels, the minimum number of experimental runs required for investigation is obtained from the relation (L-1)P + 1, where L is the number of levels and P is the number of input parameters. In the present study, since L = 3 and P = 4. Therefore minimum number of experiments need to be performed is 9. The experimental results are given in Table 3.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table 3: Experimental Layout

S. No	Inputs			
	Fuel Mixture (Butane + Pentane)	Evaporation bubble point temperature	Condensation dew point temperature	Expander inlet temperature
1.	1	1	1	1
2.	1	2	2	2
3.	1	3	3	3
4.	2	1	2	3
5.	2	2	3	1
6.	2	3	1	2
7.	3	1	3	2
8.	3	2	1	3
9.	3	3	2	1

3. Parametric estimation of the output based on the selected design

Refer section 5

Table 4: Experimental results

S. No	Output			
	First Law efficency	Exergetic efficiency	Exergy destruction	volume flow ratio
1.	11.06	51.94	1850.4	5.4133
2.	10.84	51	1921.2	5.1562
3.	10.68	50.2	1992.9	3.7904
4.	11.53	54.35	1680.1	4.8666
5.	9.33	44.37	2507.5	4.8295
6.	11.66	55.14	1627.2	7.5331
7.	10.09	47.96	2170.3	3.3212
8.	12.3	58.2	1436.5	5.0632
9.	10.14	48.51	2122.6	5.0604

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

4. S/N ratios for the experimental data were estimated

In this study lower the better is used for Exergy destruction and the ratio of specific volume flow ratio whereas higher the better is implemented for First Law efficiency and Exergetic efficiency respectively, and is given in Table 5.

Table 5:The sequences of S/N ratio

S.No	First Law efficiency	Exergetic efficiency	Exergy destruction	volume flow ratio
1.	20.8751	34.31004	-65.3453	-14.6692
2.	20.70059	34.1514	-65.6715	-14.2466
3.	20.57143	34.01407	-65.9897	-11.5737
4.	21.23659	34.70399	-64.5067	-13.7445
5.	19.39763	32.94179	-67.9848	-13.678
6.	21.33397	34.82934	-64.2288	-17.5395
7.	20.07782	33.61758	-66.7304	-10.4259
8.	21.7981	35.29846	-63.1461	-14.0885
9.	20.12076	33.71663	-66.5374	-14.0837

5. Normalized the S/N ratio

A liner normalization of the S/N ratio is performed in the range of 0 and 1, based on the formula described in section 3.2.2. Table 6 illustrates the value of normalized S/N ration, given below.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table 6 Normalized S/N ratio

S.No	First Law efficiency	Exergetic efficiency	Exergy destruction	volume flow ratio
1	0.615492	0.580586	0.545498	0.403495
2	0.542794	0.513271	0.478083	0.462902
3	0.488988	0.454998	0.412321	0.838647
4	0.766083	0.74775	0.718809	0.533485
5	0	0	0	0.542833
6	0.80665	0.800939	0.776242	0
7	0.283357	0.286756	0.259243	1
8	1	1	1	0.485127
9	0.301245	0.328786	0.29913	0.485802

6. Estimation of grey relation coefficient

The elements of the array for multiple performance characterstic listed in Table 7 represent the grey relational coefficient of each performance measure. These data were used to evaluate the correlation coefficient matrix and to determine the corresponding eigenvalues.

Table 7 Calculated grey relation coefficient

S.No	First Law efficiency	Exergetic efficiency	Exergy destruction	volume flow ratio
1	0.565286	0.543825	0.523833	0.455994
2	0.522354	0.506725	0.489277	0.482115
3	0.494554	0.478468	0.459694	0.756026
4	0.681276	0.664673	0.640048	0.517323
5	0.333333	0.333333	0.333333	0.522375
6	0.721137	0.715245	0.690839	0.333333
7	0.410967	0.412118	0.40298	1
8	1	1	1	0.492672
9	0.417099	0.426907	0.416365	0.493

7. Application of PCA to estimate the weight and thereby the grey relational grades

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

PCA may used to determine the corresponding weighting values for each performance characteristic to reflect its relative importance in the grey relational analysis. The Eigen values and explained variation for principal component is given in Table 8.

Table 8 Eigen values and explained variation for principal components

Principal component	Eigen value	Explained variation
First	3.2073	0.801804955
Second	0.7885	0.197120072
Third	0.004	0.000999975
Fourth	0.0003	7.49981E-05

The Eigen vector corresponding to each Eigen value is listed in Table 9. Moreover, the variance contribution for the first principal component characterizing the four performance characteristic is as high as 80.18%. Hence, in this research work, the squares of its corresponding eigenvectors were selection as the weighing values of the related performance characteristic. The contribution of each individual quality characteristic for principal component is shown in Table 10.

Table 9 Eigenvectors for principal component

Characteristics	Eigenvector			
	First principal	Second Principal	Third Principal	Fourth
First Law efficiency	0.552	0.162	0.727	0.374
Exergetic efficiency	0.552	0.169	0.037	0.816
Exergy destruction	0.551	0.181	0.685	0.441
volume flow ratio	0.296	0.955	0.013	0.003

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table 10 Contribution of each individual quality characteristic for principal component

Characteristic	Contribution
First Law efficiency	0.304704
Exergetic efficiency	0.304704
Exergy destruction	0.303601
volume flow ratio	0.087616

Based on the data listed in Table 7, the grey relational grades were calculated as shown in Table 11. Thus, the optimization design was performed with respect to a single grey relational grade rather than a complicated performance characteristics.

Table 11 Grey relation grade and its order

No.	Grey Relation grade	Order
1	0.536936	4
2	0.504348	5
3	0.502286	6
4	0.649757	3
5	0.350104	9
6	0.676611	2
7	0.460759	7
8	0.956168	1
9	0.426774	8

The response table of Taguchi method was employed to calculate the average grey relational grade for each process parameters of a solar driven Rankine cycle level. From the reponse table for the grey relational grades shown in Table 12, the best combination of process parameters is the set with A3 (the best fuel mixture in context to optimal performance is 0.9% Butane and 0.1% Pentane by weight), B2(Evaporation bubble point temperature=358K), C1(Condensation dew point temperature=300K), D3(Expander inlet temperature=370K).

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table 12 Response Table for grey relation grade

Symbol	Parameters	Level 1	Level 2	Level 3	Max-Min
A	Fuel Mixture (Butane + Pentane)	0.514523	0.558824	0.614567	0.100044
B	Evaporation bubble point temperature	0.549151	0.60354	0.535224	0.068316
C	Condensation dew point temperature	0.723238	0.52696	0.437716	0.285522
D	Expander inlet temperature	0.437938	0.547239	0.702737	0.264799

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

8. Apply ANOVA

ANOVA is a statistics based decision-making tool for detecting any differences in the average performance of groups of items tested. ANOVA helps in finding the significance of control factors and their interactions with the response variables by comparing the mean square against experimental errors at different confidence levels. In the present study, ANOVA was performed at a significance level of 5% i.e. $\alpha = 5\%$. The ANOVA results for the experimental data are given in Table 13.

In addition, the F test was also used to determine which parameter had a significant effect on the performance characteristic. The change is significant, when the F value is large, the results are listed in Table 13. It shows that Condensation dew point temperature is the most significant process parameter of the said cycle and it's affect the multiple performance characteristics due its highest percentage contribution amongst the process parameter. Evaporation bubble point temperature has the least effect on the performance of the solar driven Rankine cycle.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Table 13 Results of the analysis of variance

Symbol	Parameters	DOF	Sum of Square	Mean square	F Value	Contributions (%)
A	Fuel Mixture (Butane + Pentane)	2	0.01508	0.007539	0.19	5.864167525
B	Evaporation bubble point temperature	2	0.007819	0.003910	0.09	3.040578639
C	Condensation dew point temperature	2	0.1280	0.06401	2.97	49.77542727
D	Expander inlet temperature	2	0.1062	0.05312	2.11	41.29804981
Error		2	0.000056	0.0005		0.021776749
Total		8	0.257155	0.129079		100

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

9. Conformity test, to check the robustness of the optimal parameters.

The results of confirmation experiments have also been presented to validate the optimal results. A confirmation experiment is the final step of a design of experiment. Its purpose is to verify that the optimum conditions suggested by the matrix experiment do indeed give the projected improvement. The confirmation experiment is performed by conducting a test with optimal settings of the factors and levels previously evaluated. The optimal values can be predicted using equation

$$\eta_0 = \eta_m + \sum_{i=1}^{j}(\eta_i - \eta_m) \quad (5.1)$$

Where, j is the number of factors that significantly affects the performance measure and η_m is the mean value of S/N ratios in all experimental runs, η_i are the S/N ratios corresponding to optimum factor levels.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

In order to validate the results obtained, confirmation experiments were conducted for each of the performance characteristics (First Law efficiency, Exergetic efficiency, Exergy destruction and volume flow ratio) at optimal levels of the process variables. The average values of confirmation results have been compared with the predicted values. The results are listed in Table 14.

Table 14: Results of Confirmation experiment

Performance characteristics	**Estimated Optimal value**	**Actual value (Value obtained through confirmation experiment)**
First Law efficiency	12.3	12.2
Exergetic efficiency	58.2	58.25
Exergy destruction	1436.5	1437.5
volume flow ratio	5.0632	5.1632

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

7 Conclusions

This paper explored the effect of four parameters of a solar driven Rankine cycle i.e. First Law efficiency, Exergetic efficiency, Exergy destruction and volume flow ratio on First Law efficiency, Exergetic efficiency, Exergy destruction and volume flow ratio. Further, grey relational analysis along with principal component analysis was used to perform multi-response optimization of a solar driven Rankine cycle in order to obtain optimal combination of the cycle parameters. In addition, ANOVA was employed to determine the significance of the controllable parameters. Based on the results of the present study, following conclusions are drawn:

- Grey relational analysis along with the principal component analysis is a simple and effective multi-response optimization method since principal component analysis provides different weighting values of the responses reflecting the relative importance for each response in the multi-response optimization problem.
- The optimal combination of process parameters is the set with A3 (the best fuel mixture in context to optimal performance is 0.9% Butane and 0.1% Pentane by weight), B2(Evaporation bubble point temperature=358K), C1(Condensation dew point temperature=300K), D3(Expander inlet temperature=370K).

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

- The major controllable parameter that significantly affects the multi-performance characteristics of the solar driven Rankine cycle is the Condensation dew point temperature.
- Confirmation experiments were conducted for each of the performance characteristics (First Law efficiency, Exergetic efficiency, Exergy destruction and volume flow ratio) at optimal levels of the process variables that validate the obtained results.
- This piece of research work will be helpful to users of solar energy, academician, researchers, and other concerned persons, in understanding the importance, severity and benefits obtained by the application, implementation and optimization of the cyclic parameters of solar driven Rankine cycle.

In this research work, the Taguchi based grey relational analysis coupled with PCA has been successfully carried out whereas, for any optimized solution, it is required to have a real time scenario that may be taken into consideration by the application of different soft computing techniques like genetic algorithm, Simulated annealing, etc.

The result generated are purely based on theoretical modelling and for further research, experimental analyses are required to consolidate the generated results.

References

Abadi, G. B., Yun, E. and Kim, K. C. (2015), "Experimental study of a 1 kw organic Rankine cycle with a zeotropic mixture of R245fa/R134a", *Energy*, Vol. 93, pp. 2363-2373.

Aghahosseini, S and Dincer. I. (2013), "Comparative performances analysis of low temperature Organic Rankine Cycle (ORC) using pure and zeotropic working fluids", *Applied Thermal Engineering*, Vol. 54 No. 1, pp. 35-42.

Andreasen, J. G., Larsen, U., Knudsen, T., Pierobon, L. and Haglind, F. (2014), "Selection and optimization of pure and mixed working fluids for low grade heat utilization using organic Rankine cycles", *Energy*, Vol. 73 , pp. 204-213.

Asjad, M., Kulkarni, M. S. and Gandhi, O. P. (2013), "Opportunistic Actions for Subassemblies of a Reciprocating Compressor: An LCC-based Approach", *International Journal of Performability Engineering*, Vol 9. No 3., pp. 273-285.

Chen, H. (2010). *The conversion of low-grade heat into power using supercritical Rankine cycles* (Doctoral dissertation, University of South Florida).

Chen, H., Goswami, D. Y., Rahman, M. M. and Stefanakos, E. K. (2011), "A supercritical Rankine cycle using zeotropic mixture working fluids for the conversion of low-grade heat into power" *Energy*, Vol 36. No. 1, pp. 549-555.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

Deethayat, T., Asanakham, A. and Kiatsiriroat, T. (2016), "Performance analysis of low temperature organic Rankine cycle with zeotropic refrigerant by Figure of Merit (FOM)". *Energy*, Vol. 96, pp. 96-102.

Deethayat, T., Kiatsiriroat, T. and Thawonngamyingsakul, C. (2015), "Performance analysis of an organic Rankine cycle with internal heat exchanger having zeotropic working fluid", *Case Studies in Thermal Engineering*, Vol. 6, pp. 155-161.

DiPippo, R. (2004), "Second law assessment of binary plants generating power from low-temperature geothermal fluids", *Geothermics*, Vol. 33 No. 5, 565-586.

Feng, Y., Hung, T., Greg, K., Zhang, Y., Li, B. and Yang, J. (2015), "Thermoeconomic comparison between pure and mixture working fluids of organic Rankine cycles (ORCs) for low temperature waste heat recovery", *Energy Conversion and Management*, Vol. 106, pp. 859-872.

Feng. Y, Hung. T., Zhang. Y., Li. B.,Yang. J. and Shi.Y. (2015), "Performance comparision of low grade ORCs(Organic Rankine cycles) using R245fa,pentane and their mixtures based on the thermoeconomic multi-objective optimization and decision making", Vol. 90, pp. 2018-2029.

Geng, D., Du, Y. and Yang, R. (2016), "Performance analysis of an organic Rankine cycle for a reverse osmosis desalination system using zeotropic mixtures", *Desalination*, Vol. 381, pp. 38-46.

Lemmon, E. W., Huber, M. L. and McLinden, M. O. (2002), "NIST reference fluid thermodynamic and transport properties–REFPROP", *NIST standard reference database*, 23, v7.

Gopalsamy, B. M., Mondal, B. and Ghosh, S. (2009), "Taguchi method and ANOVA: An approach for process parameters optimization of hard machining while machining hardened steel", *Journal of Scientific and Industrial research,* Vol. 68 No. 8, pp. 686-695.

Guo, C., Du, X., Yang, L. and Yang, Y. (2015), "Organic Rankine cycle for power recovery of exhaust flue gas", *Applied Thermal Engineering*, Vol. 75, pp. 135-144.

Herberle, F., Preibinger, M., and Bruggemann,D. (2012), "Zeotropic mixtures as working fluid in organic Rankine cycles for low-enthalpy geothermal resources", *Renewable Energy* Vol. 37, pp. 364-370.

Herberle, F., Schiffechner, C. and Bruggemann, D. (2016), "Life cycle assessment of Organic Rankine cycle for geothermal power generation considering low-GWP working fluids", *Geothermic*, Vol. 64, pp. 392-400.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

© Emerald Publishing Limited

This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Imran, M., Usman, M., Park, B. S. and Yang, Y. (2016), "Comparative assessment of Organic Rankine Cycle integration for low temperature geothermal heat source applications", *Energy*, Vol. 102, pp. 473-490.

Jolliffe, I. (2002). *Principal component analysis*. John Wiley & Sons, Ltd.

Kang, Z., Zhu, J., Lu, X., Li, T. and Wu, X. (2015), "Parametric optimization and performance analysis of zeotropic mixtures for an organic Rankine cycle driven by low-medium temperature geothermal fluids", *Applied Thermal Engineering*, Vol. 89, pp. 323-331.

Khan, S., Asjad, M., and Ahmad, A. (2015), "Review of Modern Optimization Techniques", *International Journal of Engineering Research and Technology*, Vol. 4 No. 04, pp. 984-988.

Kumar, U., Karimi, M. N. and Asjad, M. (2016), "Parametric optimisation of the organic Rankine cycle for power generation from low-grade waste heat", *International Journal of Sustainable Energy*, Vol. 35 No. 8, pp. 774-792.

Lee, H. Y., Park, S. H. and Kim, K. H. (2016), "Comparative analysis of thermodynamic performance and optimization of organic flash cycle (OFC) and organic Rankine cycle (ORC)", *Applied Thermal Engineering*, Vol. 100, PP. 680-690.

Li, Y. R., Du, M. T., Wu, C. M., Wu, S. Y. and Liu, C. (2014), "Potential of organic Rankine cycle using zeotropic mixtures as working fluids for waste heat recovery", *Energy*, Vol. 77, pp. 509-519.

Lu, J., Zhang, J., Chen, S. and Pu, Y. (2016), "Analysis of organic Rankine cycles using zeotropic mixtures as working fluids under different restrictive conditions", *Energy Conversion and Management*, Vol. 126, pp. 704-716.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Vol. 2 No. 11, pp. 559-572.

Peris, B., Navarro-Esbrí, J., Molés, F., Martí, J. P. and Mota-Babiloni, A. (2015), "Experimental characterization of an Organic Rankine Cycle (ORC) for micro-scale CHP applications", *Applied Thermal Engineering*, Vol. 79, pp. 1-8.

Sadeghi, M., Nemati, A. and Yari, M. (2016), "Thermodynamic analysis and multi-objective optimization of various ORC (organic Rankine cycle) configurations using zeotropic mixtures", *Energy*, Vol. 109, pp.791-802.

Schoenmaker, J., Rey, J. F. Q. and Pirota, K. R. (2011), "Buoyancy organic Rankine cycle", *Renewable energy*, Vol. 36 No. 3, pp. 999-1002.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

Selvaraj, D. P. and Chandramohan, P. (2010), "Optimization of surface roughness of AISI 304 austenitic stainless steel in dry turning operation using Taguchi design method", *Journal of engineering science and technology*, Vol. *5 No.* 3, pp. 293-301.

Shu, G., Gao, Y., Tian, H., Wei, H., and Liang, X. (2014), "Study of mixtures based on hydrocarbons used in ORC (Organic Rankine Cycle) for engine waste heat recovery", *Energy*, Vol. 74, pp. 428-438.

Tchanche, B. F., Papadakis, G., Lambrinos, G. and Frangoudakis, A. (2009), "Fluid selection for a low-temperature solar organic Rankine cycle", *Applied Thermal Engineering*, Vol. 29 No. 11, pp. 2468-2476.

Wang, M., Wang, J., Zhao, Y., Zhao, P. and Dai, Y. (2013), "Thermodynamic analysis and optimization of a solar-driven regenerative organic Rankine cycle (ORC) based on flat-plate solar collectors", *Applied Thermal Engineering*, Vol. *50 No.* 1, pp. 816-825.

Wu Y, Zhu Y, and Yu Li. (2016), "Thermal and economic performance analysis of zeotropic mixtures for Organic Rankine Cycles", *Applied Thermal Engineering*, Vol. 96, pp. 57-63.

Xu, G., Song, G., Zhu, X., Gao, W., Li, H. and Quan, Y. (2015), "Performance evaluation of a direct vapor generation supercritical ORC system driven by linear Fresnel reflector solar concentrator", *Applied Thermal Engineering*, Vol. 80, pp. 196-204.

Yamada, N., Minami, T. and Mohamad, M. N. A. (2011), " Fundamental experiment of pumpless Rankine-type cycle for low-temperature heat recovery", *Energy*, Vol. 36 No. 2, pp. 1010-1017.

Yamada, N., Watanabe, M. and Hoshi, A. (2013), "Experiment on pumpless Rankine-type cycle with scroll expander", *Energy*, Vol. 49, pp. 137-145.

Yamamoto, T., Furuhata, T., Arai, N. and Mori, K. (2001), "Design and testing of the organic Rankine cycle", *Energy*, Vol. 26 No. 3, pp. 239-251.

Zhai, H., An, Q., Shi, L., Lemort, V. and Quoilin, S. (2016), "Categorization and analysis of heat sources for organic Rankine cycle systems", *Renewable and Sustainable Energy Reviews*, Vol. 64, pp. 790-805.

Zhao, L. and Bao, J. (2014), "Thermodynamic analysis of organic Rankine cycle using zeotropic mixtures", *Applied Energy*, Vol. 130, pp. 748-756.

Zhou, Y., Zhang, F. and Yu, L. (2016), "Performance analysis of the partial evaporating organic Rankine cycle (PEORC) using zeotropic mixtures", *Energy Conversion and Management*, Vol. 129, pp. 89-99.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

Biographical notes:

Deepak Tiwari is pursuing Ph.D. From Department of Mechanical Engineering of Jamia Millia Islamia University, Delhi, India. His research area includes solar thermal power generation.

Dr. A.F.Sherwani is working as an Associate Professor in Department of Mechanical Engineering of Jamia Millia Islamia University, Delhi. His research area includes Solar Energy , alternate refrigeration technologies, exergy analysis of thermal systems and renewable, sustainable energy and life cycle analysis. He has published plenty of research paper in reputed international and national journals. He has worked as reviewer in several reputed general.

Dr Mohammad Asjad is currently an Assistant Professor with the Discipline of Mechanical Engineering at Jamia Millia Islamia, New Delhi. He graduated in Mechanical Engineering from Z.H.C.E.T (AMU, Aligarh) and later did his Masters in Industrial and Production Engineering from Z.H.C.E.T (AMU, Aligarh) and securing first position in both of them. Subsequently, he completed his Ph.D. from Indian Institute of Technology Delhi (IITD). He also receipts the scholarships during Master and Research at A.M.U and IIT Delhi, respectively. His research interests are: Strategic Asset Management, Reliability Engineering, Maintenance Management, Manufacturing Operations Planning, Life Cycle Costing, Soft computing and Applied Artificial Intelligence. He has published several research articles at various national and international journals and conferences of repute. He is also on the panel of Editorial board and reviewer for various reputed international journals and conferences.

Dr. Akhilesh Arora is an Assistant Professor in the Department of Mechanical Engineering at Delhi Technological University (erstwhile Delhi College of Engineering) Delhi since 2004. Prior to joining Delhi College of Engineering, he worked at College of Military Engineering, Pune in the capacity of Lecturer for two years. He has also served as Associate professor and Head of Mechanical and Automation Engineering Department at Indira Gandhi Delhi Technical University of Women, Delhi during the period from 2012 to 2014. He obtained his Ph.D. and Master's Degree from the Indian Institute of Technology, Delhi in the year 2010 and 1997 respectively. His research areas are Refrigeration and Air Conditioning, Thermal Science and Engineering; Energy Conservation and Waste Heat Recovery. He has published number of research papers in International and national journals. Dr. Arora is an active member of ASHRAE and ISHRAE.

© Emerald Publishing Limited
This is a pre-print of a paper and is subject to change before publication. This pre-print is made available with the understanding that it will not be reproduced or stored in a retrieval system without the permission of Emerald Publishing Limited.

Downloaded by Delhi Technological University At 21:22 02 July 2017 (PT)

2017 9th International Conference on Communication Systems and Networks (COMSNETS)

Group Authorization using Threshold Signatures for Medical Procedures

Preeti Nagar, Divyashikha Sethia
Department of Computer Engineering
Delhi Technological University, New Delhi, India
preeti.nagar93@gmail.com, divyashikha@dce.edu

***Abstract*—Digitalisation of the health sector is being adopted by many countries to deliver reliable healthcare services.In healthcare organisations, frauds are rampant possessing a serious threat to good governance.To control such malpractices we present the design of a system for group authorization in healthcare organisations.In this medical professionals are provided with some shares of a secret.To authorise any medical document at least threshold number of medical professionals have to offer their signatures shares.Only when enough number of valid signature shares get combined, the document gets authorised.Maintaining audit logs using the group signature can ensure non-repudiation.We implement threshold Gap Diffie-Hellman (GDH) signature scheme based on the concept of secret sharing for the proposed system.Also, we compare the results of our implementation with the basic RSA scheme.Our implementation has shorter key generation time as compared to a serial RSA scheme.Also, unlike RSA scheme supports the very idea of threshold signatures.Incorporating proposed group authorization system can prevent frauds and corruption in the healthcare sector by providing transparency and accountability.**

I. Introduction

Consider the scenario of the health sector.Widespread corruption is the undeniable reality of healthcare.The problems are actually of a multidimensional nature.They involve frauds from both ends, from the side of medical officials as well as from the side of people seeking services.Medical officials are often involved in scams related to the supply of medicines and goods, organ transplantation and even purchase of instruments.Patients are involved in over billing of insurance claims and fraud reimbursements of medical bills.People are entitled to a lot of medical benefits primarily public sector beneficiaries.In developing countries such as India, government employees and pensioners and their dependents get comprehensive health care facilities [1] where medical malpractices may be used to seek reimbursement for false medical bills or prescribe costly medicines and procedures which may not be required.In European countries such as Denmark, public funds finance around 85% of health care expenses, thus health care is to a considerable extent a public task [2]. Many times people make up false scenarios and manage to get fraud reimbursements.The Czech Republic, Slovakia, Romania, Italy, Bulgaria and Greece are considered having a widespread corruption problem and seem to encounter more frauds in medical service delivery and misuse of (high) level positions [3].If approval from a certain number of authorities is made mandatory for authorising any document, then it will make it difficult for frauds to happen quickly.That is where threshold signatures can play a significant role.

Digital signature is for authorization for digital documents or data. Authorization means the receiver believes that a known sender created the message.The idea of threshold signature is to distribute a secret key among some participants (say n) and the signature can be generated when a subset of t (threshold) or more players combine their shares of the secret key.Anyone using public keys can verify the signature made.

In this paper, we present design of a system for group authorization of medical documents in healthcare organisations.For this we implement a threshold Gap Diffie-Hellman (GDH) signature scheme proposed in [4] as a part of this work.It is based on the GDH signature scheme [5].Shares of the secret key are provided to the participants (here which would include medical officials)on cards.A medical official can be a doctor or an administrative official whose approval is required in the hierarchy.Whenever some document needs to be authorised at least t(threshold) number of medical officials need to give their authorization using their cards and only when at least t valid signature shares are combined can the document be digitally signed and hence approved.The factor which is important here is the value of the threshold for signature generation protocol.If the value of threshold t is equal to n (number of participants), then the signature shares from all the members are needed for a valid authorization.If the value of the threshold is 1, then authorization one participant is enough to authorise a document.Our scheme of getting digital signatures is similar to getting signatures of a certain number of authorities on paper for approval.

The rest of the paper is organised as follows.Section II reviews the necessary background.Section III discusses related work in threshold signatures.Section IV presents the design and algorithm of the proposed system. Section V discusses implementation and results.Finally, we conclude in section VI.

II. Background

In this section, we briefly describe some background information needed to understand the proposed design of group authorization system and our implementation of the threshold signature scheme.

Secret sharing is the concept of distributing a secret value amongst some players in the form of parts referred to as shares, and at least t (t is the threshold) number of these shares

978-1-5090-4250-0/17/$31.00 ©2017 IEEE

have to be combined to access the secret value.Secret sharing was invented independently by Adi Shamir [6] and George Blakley [7] in 1979.Shamir's scheme is more space-efficient than Blakley's as in Shamir's scheme the size of the shares is only as large as the original secret while in Blakley's scheme shares are t times greater, where t is the threshold number of players.

Shamir [6] states the problem as described below.Let D be some secret data which we divide into n pieces D_1, $D_2 \cdots D_n$ in such a way that:

(1) Computation of D is straightforward if any k or more D_i pieces are known;

(2) Deduction of D is completely undetermined if k - 1 or fewer D_i pieces are known(in the sense that reveals no information on D).

Such a scheme is called a (k, n) threshold scheme.We have used Shamir's [6] scheme in our implementation of the threshold signature scheme.

A. Digital Signature Schemes

Many digital signature schemes have evolved during last few decades. RSA signature scheme [8] uses RSA cryptosystem for signing and verifying a message. The digital signature scheme changes the roles of public and private keys.ElGamal digital signature scheme [9] uses the same keys as RSA signature scheme [8] but the signature generation and verification algorithm are different.To reduce the size of the signature, Schnorr a proposed a new scheme [10] based on ElGamal but with a reduced signature size.

There are several variations and additions to the central concept of digital signatures.We briefly discuss some of them here.

• Time Stamped Signatures: Sometimes we need to associate a timestamp with a signed document to prevent it from being replayed by a foe. This method called time-stamped digital signature scheme.

• Blind Signatures: Sometimes we have a document that we want to get signed without disclosing the data in the document to the signer. For example, a researcher might have discovered a critical theory that needs to be signed by a public authority without allowing the other person to know the contents of the method. Such scheme is referred to as blind signature scheme.

• Threshold Signatures: In threshold signature schemes parts of a secret are distributed among all the participants (say n).To sign a message M any set of members of size at least t, where t is the threshold value p can use their shares of the secret and generate a signature of M that can be verified by anyone using the unique fixed public key.

B. Bilinear Maps and GDH Groups

Let G1, G2 and GT be multiplicative cyclic groups of prime order p and g1 and g2 be generators of G1 and G2 respectively.A bilinear map e from G1 $\times$ G2 to GT is a function e : G1 $\times$ G2 $\rightarrow$ GT for u $\in$ G1 , v $\in$ G2 and a, b $\in Z_p$ .The bilinear map e satisfies the following properties:

1. e(u^a , v^b) = e(u, v)^{ab} and

2. e (g1, g2) $\neq$ 1.

If G1 == G2, it is called a symmetric pairing, otherwise the pairing is asymmetric. Bilinear maps associate pairs of elements from G1 and G2 with elements in GT and thus, also referred to as pairings.

Consider two problems in G.

(1) Decisional Diffie-Hellman (DDH) problem: Given four group elements (g, u, v, h) determine whether they satisfy the relation $\log_g u = \log_v h$ or are random.If the four elements satisfy this property, they are referred to as a valid Diffie-Hellman (DH) tuple.

(2) Computational Diffie-Hellman (CDH) problem: Given three random group elements compute h = $g^{log_g u.log_g v}$

A prime order group G is called a gap Diffie-Hellman (GDH) group if CDH is hard in G but DDH is easy. Computation of DSS signatures is faster than the computation of RSA signatures when using the same p.

C. GDH Signature Scheme

The GDH signature scheme [5] works in GDH groups.The signature scheme consists of three algorithms.The key generation algorithm (K) uses some global information I, picks random secret key x and computes public key v corresponding to the secret key x.The signature generation algorithm (S) takes a secret key x, a message M to sign and global information I.It computes hash of the message and then generates a signature σ for the message M.Given a public key v, a message M and a signature σ the verification algorithm (V) checks that if the signature is valid or not.It checks for valid Diffie-Hellman (DH) tuple property in this step.

III. Related Work

There are many threshold signature schemes proposed in past.RSA based threshold signature scheme [11] is one of them.Shoup's scheme [11] requires a trusted dealer to generate keys while in the GDH threshold signature scheme, keys can be generated with or without a trusted dealer.Proactivization means to renew the shares of a secret periodically so that if someone is able to get some shares, less than the threshold value one time, then no attack is possible by getting more number of shares next time when all shares have been renewed.Proactivization is not discussed in RSA threshold scheme while GDH scheme [4] can be proactivized using the proactive secret sharing algorithm [12].

We compare our implementation with primary RSA signatures [8].RSA signatures can be utilised for authorization but unlike are our implementation does not provide a means to authorise any document only when at-least threshold number of medical professionals combine their signature shares.Also, the key generation overhead in RSA is high as we need public-private key pairs for all the participants, unlike our implementation in which only one public key is required and private key shares are generated for each participant.We prove these arguments in our implementation and result section by comparing the timings for both the schemes.

IV. Threshold GDH Signature Scheme

The threshold GDH signature scheme [4] is the threshold variant of the GDH signature scheme [5].Let G be a GDH group with the generator as g and of prime order p. Let H $\{0,1\}^* \rightarrow G^*$ be a hash function which maps arbitrary long strings to G*.The global information I contain (g, p, H).

The threshold signature scheme TS = (TK, TS, V) consists of three algorithms: the threshold key generation algorithm (TK), threshold signature generation algorithm TS and the verification algorithm V.The verification algorithm V in TS is same as of GDH signature scheme of [5].

The threshold key generation algorithm (TK) takes global information I and outputs the public key pk and the secret key x_i of each player such that such that

$$(x_1, x_2, \cdots x_n) \xrightarrow{(t,n)} (sk) \tag{1}$$

where sk is the secret key associated with the public key pk.

The threshold signature generation algorithm consists of two parts: signature share generation and signature reconstruction. In signature share generation each player calculates its signature share.In signature reconstruction, the signature shares are verified to identify the honest players and their shares are combined.The output of the algorithm is the signature σ corresponding to the message M.

In signature verification, a verifier just checks whether (g, pk, H(M),σ) is a valid Diffie-Hellman tuple to verify the validity of a signature σ of a message M.

A. Design and Algorithm

In this section, we present the design of our system for group authorization in heath care organisations.We then describe the algorithm for implementing the same using threshold signatures.

Consider the scenario of a hospital.When any document D needs to be authorised in any situation such as one in which a beneficiary is to be authorised to get reimbursement for medicines prescribed to her or in case a decision for purchase of expensive instruments is to be made, the document is generated digitally with all the relevant details.Then the medical officers provide their signatures for authorising that document using their cards provided to them having their signature share.Only when a sufficient number of medical officials (at least threshold) provide their signature is the document authorised otherwise it not considered to be a valid one.Also, we maintain audit logs so as to know at later stages which all physicians verified a particular document.Thus, we achieve non-repudiation in sense nobody can deny authorising a particular document.Fig.1 shows the design of the system.We use the terms medical officer, practitioners, doctors, medical official or simply official interchangeably all implying the same meaning in our context.

We now describe the details of the algorithm for implementing such a system.In the scheme implemented there are N medical practitioners or doctors (D1, D2, D3,.... Dn) and a server. Table I gives the various symbols used in the scheme. Fig. 2 shows the sequence diagram for the

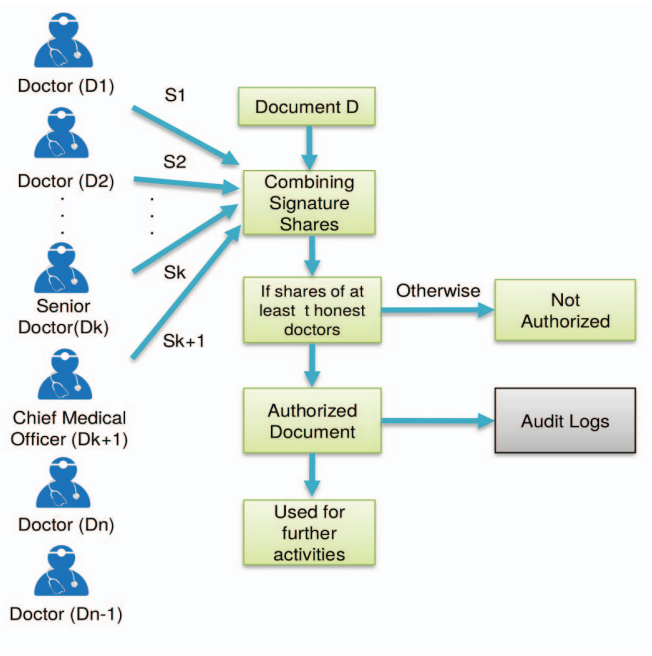


Fig. 1. System Design.

implemented scheme. The steps of the algorithm are described below

1. Server generates secret key (sk) and public key (pk) such that

$$pk = g^{sk} \tag{2}$$

where, g generator of G.

2. Server forms the polynomial P(x) such that degree (P(x)) = t and P(0) = sk.

$$P(x) = sk + ax + bx^2 + cx^3 + \cdots hx^t \tag{3}$$

3. Then initializing id of medical officials to be id_1, id_2, ... id_n and generates key shares for each of them as

$$k_i = P(id_i) \tag{4}$$

and the sends the id_i and the corresponding k_i to the official i via secure channel.

4. The server then sends the message M via the broadcast channel to all the N medical officials.

5. Each official calculates the signature share S_i using the k_i and the message M received from the server

$$S_i = H(M)^{k_i} \tag{5}$$

6. Then the officials who want to form group signature (Set R) send their corresponding S_i to the server via a secure channel.

TABLE I
SYMBOLS USED IN THE SCHEME

Symbol	Meaning
N	Number of medical officials
T	Threshold
id_i	Identifier for a Medical official
Sk	Secret key
Pk	Public key
K_i	Key share of medical official i
M	Message
H(M)	Hash Function
S_i	Signature share of medical official i
L_i	Lagrange coefficient for medical official i
S	Signature generated

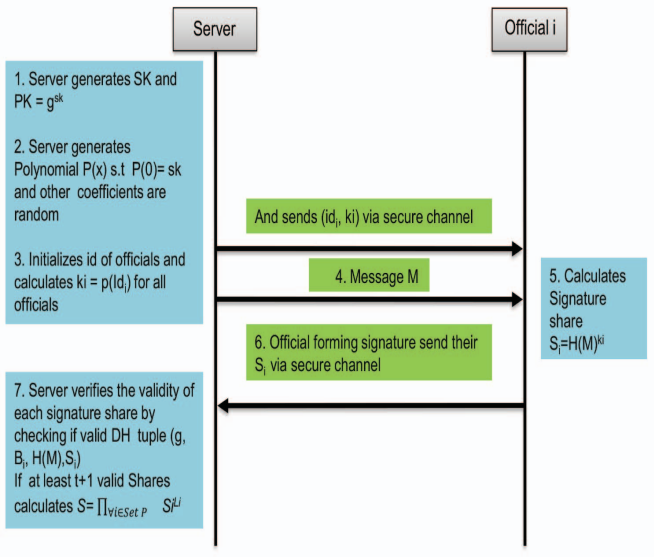


Fig. 2. Sequence diagram for the scheme.

7. The server then verifies the validity of signature share S_i for each official of set R.The server for this calculates

$$B_i = g^{k_i} \tag{6}$$

And then simply checks whether (g, B_i, H(M), S_i) is a valid Diffie-Hellman tuple for each official i of set R. If this does not hold for some official, that official is assumed malicious. If the server finds at least t+1 honest officials it computes the resulting signature otherwise throws an error message invalid signature shares, signature generation unsuccessful. Now in a case of at least t+1 honest officials (lets say set Q of w fair officials such that w≥t+1)the server computes Lagrange coefficient (L_i) for each physician i of set Q.

Let us assume medical official in set Q have id's as id_1, id_2, id_3,... id_w.The value of L_i

$$L_i = \prod_{\forall j \in SetQ, j \neq i} \frac{id_j}{id_j - id_i} \tag{7}$$

For example for official with id_1 the L_1 will be

$$\frac{id_2}{id_2 - id_1} \times \frac{id_3}{id_3 - id_1} \times \cdots \frac{id_w}{id_w - id_1} \tag{8}$$

The server then computes the signature S as:

$$S = \prod_{\forall i \in SetQ} S_i^{L_i} \tag{9}$$

i.e if we take

$$D = \sum_{\forall i \in SetQ} (K_i \times L_i) \tag{10}$$

where i is official in set Q and then

$$S = H(M)^D \tag{11}$$

is the resulting signature.
Explanation: if K_i and L_i are valid and correctly calculated the by using the concept of polynomial interpolation using Lagranges method D will be equal to the P(0) i.e sk here.

8. The signature generated can be verified by just checking whether (g, pk, H(M), S) is a valid Diffie-Hellman tuple or not.

In the following section, we look into the results obtained corresponding to our implementation of the algorithm described above.

V. IMPLEMENTATION AND RESULTS

Our implementation of Threshold GDH signature scheme [4] is built upon Pairing-Based Cryptography (PBC) library [13].The implementation results are shown below.We analyse the results obtained for our implementation and also compare them with the basic RSA signature scheme.The time in all the graphs is taken to be in seconds.

The graphs corresponding to the to the timings recorded are shown in the following section.Fig. 3 indicates the plot of key generation time (includes the public key, secret key and secret key shares for medical officials)versus the number of medical officials.We observe that the key generation time increases gradually with the increase in the number of medical officials.

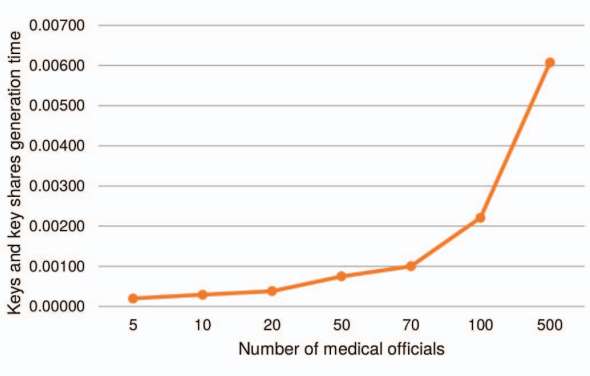


Fig. 3. Key and key shares generation time versus number of medical officials plot.

Fig. 4 presents the plot of signature share generation time versus the number of officials.Fig. 5 represents the plot of the total time of the scheme for the threshold value of 4 (random) versus the number of officials.Total time here includes signature

generation time along with keys generation step and signature share generation steps.

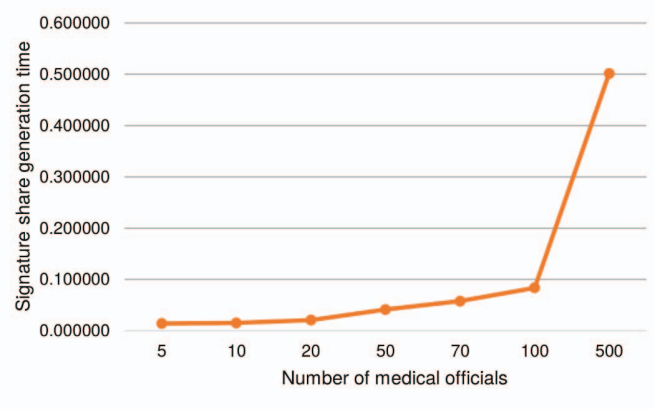


Fig. 4. Signature shares generation time versus number of medical officials plot.

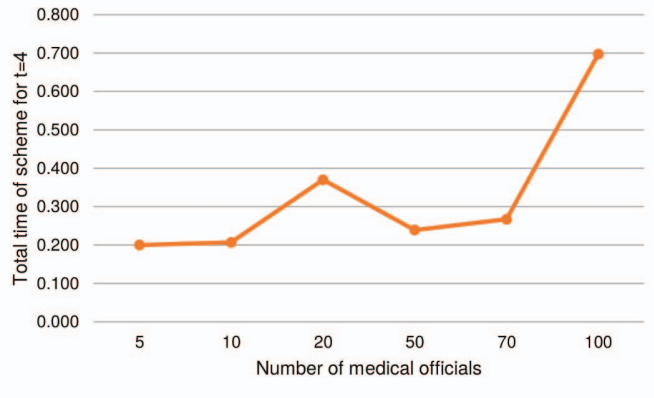


Fig. 5. Total time for signature generation (for t=4) versus number of medical officials.

Fig. 6 shows the plot for the timings obtained for the signature generation corresponding to the different values of the number of officials combining their signature shares to generate the signature.

For some real-time situations, medical documents might require authorization from officials across organisations.This would lead to increase in the number of officials for whom key shares are generated under a single authority.But, managing the authority in hierarchical form would result in easier tracking of the corrupt officials in comparison to keeping a vast number of officials under a single authority.Considering this fact, the number of officials in the results are taken to be in the range of hundreds.

1) Comparison with serial RSA scheme: We now compare our implementation of GDH scheme with the basic RSA scheme implementation.RSA scheme is suitable for basic authorization as it does not provide a means to ensure authorization of any document only when a threshold number of officials provide their signature shares.We consider the key

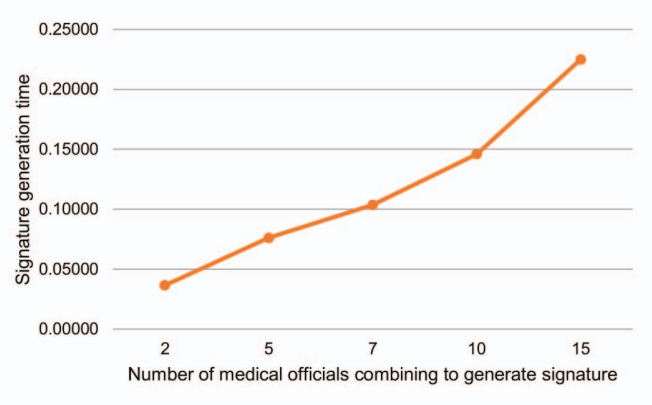


Fig. 6. Signature generation time versus number of medical officials combining signature shares plot.

generation timings in the RSA scheme [14] for a different number of medical officials.Fig. 7 shows the plot of key generation time for the RSA scheme and our implementation.We clearly observe that the RSA scheme has a high key generation overhead in comparison to our implemented scheme.Thus, our implementation of threshold signature proves to be better for the proposed system for group authorization of medical documents.

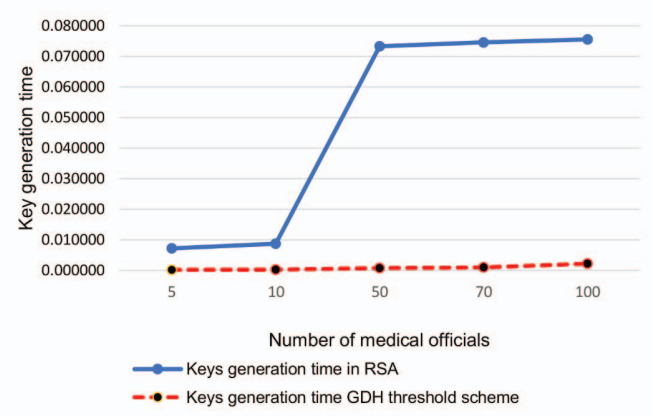


Fig. 7. Key Generation Time for implemented threshold signature scheme and basic RSA scheme.

VI. Conclusion

In this paper, we presented the design a system for group authorization of documents in heath sector.In this system, to authorise a document valid signature shares from at least a threshold number of medical officials have to be combined.For this, we successfully implemented the threshold Gap-Diffie Hellman signature scheme.Our implementation of the scheme is based on the Pairing-based cryptographic library.We have shown the results obtained for our implementation and have also compared them with the basic RSA scheme.The comparison shows that our implemented scheme has lesser key generation overhead and also supports the idea of threshold signatures suitable for our proposed system.Our system for group authorization

using the threshold signature scheme implemented can prove to be useful in preventing frauds happening in health sector worldwide.

REFERENCES

[1] Central government health scheme,government of india. [Online]. Available: http://msotransparent.nic.in/cghsnew/index.asp

[2] (2008) Healthcare in denmark ministry of health slotsholmsgade 10-12 dk-1216 copenhagen k. [Online]. Available: http://www.ilo.org/dyn/travail/docs/2047/health/

[3] (2013, Oct.) Study on corruption in the healthcare sector,european commission directorate-general home affairs. [Online]. Available: http://ec.europa.eu/dgs/home-affairs/what-is-new/news/news/2013/20131219_01_en.htm

[4] A. Boldyreva, "Threshold signatures, multisignatures and blind signatures based on the gap-diffie-hellman-group signature scheme," in *International Workshop on Public Key Cryptography*. Springer, 2003, pp. 31–46.

[5] D. Boneh, B. Lynn, and H. Shacham, "Short signatures from the weil pairing," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2001, pp. 514–532.

[6] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[7] G. R. Blakley, "Safeguarding cryptographic keys," *Proc. of the National Computer Conference1979*, vol. 48, pp. 313–317, 1979.

[8] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[9] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," in *Workshop on the Theory and Application of Cryptographic Techniques*. Springer, 1984, pp. 10–18.

[10] C.-P. Schnorr, "Efficient signature generation by smart cards," *Journal of cryptology*, vol. 4, no. 3, pp. 161–174, 1991.

[11] V. Shoup, "Practical threshold signatures," in *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 2000, pp. 207–220.

[12] A. Herzberg, M. Jakobsson, S. Jarecki, H. Krawczyk, and M. Yung, "Proactive public key and signature systems," in *Proceedings of the 4th ACM conference on Computer and communications security*. ACM, 1997, pp. 100–110.

[13] B. Lynn. The pairing-based cryptography (pbc) library. [Online]. Available: http://crypto.stanford.edu/pbc

[14] A. Kiluk. (2013, Dec.) Rsa-library. [Online]. Available: https://github.com/andrewkiluk/RSA-Library

Holographic Dark Energy Model in Brans-Dicke Theory with Future Event Horizon

Milan Srivastava [2] **and C. P. Singh** [1]

[1,2]Department of Applied Mathematics,
Delhi Technological University
Bawana Road, Delhi-110 042, India.
[1]cpsphd@rediffmail.com
[2]milandtu@gmail.com

arXiv:1706.06777v1 [gr-qc] 21 Jun 2017

Abstract In this paper, we study the dynamics of non-interacting and interacting holographic dark energy models in the framework of Brans-Dicke theory. As system's infra-red cut-off we consider the future event horizon. The motivation of this work is to use the logarithmic form of the Brans-Dicke scalar field, $\phi \propto ln(\alpha + \beta a)$, where α and β are constants and 'a' is the scalar factor as proposed in a recent work [29] to study the new agegraphic dark energy models. We find the time-dependent equation of state parameter and deceleration parameter which describe the phase transition of the universe. We observe that the model explains the early time inflation and late time acceleration including matter-dominated phase. It is also observed that the equation of state parameter may cross phantom divide line in late time evolution. The cosmic coincidence problem is also discussed for both the models. We observe that this logarithmic form of Brans-Dicke scalar field is more appropriate to achieve a less acute coincidence problem in non-interacting model whereas a soft coincidence can be achieved if coupling parameter in interacting model has small value.

PACS NO: 04.20.-q; 04.20.-Jb.
Keywords- Cosmology; FRW model; Holographic dark energy; Brans-Dicke theory; Coincidence problem.

[1]Corresponding author

1 Introduction

The astrophysical observations [1–5] have led to many interesting phenomena in the field of cosmology, especially, the accelerating expansion of the Universe. This event has brought many challenges to the cosmologists. The observations show that the accelerating expansion of the Universe is due to a mysterious form of energy with negative pressure which is dubbed as "dark energy" (DE). Nowadays, the DE has become one of the most active field in physics and astronomy. Many models have been proposed to describe this late -time acceleration of the Universe. The simplest candidate for DE is the cosmological constant [6]. However, it suffers the so-called cosmological constant (CC) problem (the fine-tuning problem) and the cosmic coincidence problem [7, 8].

In order to alleviate the CC problems and explain the accelerated expansion, some alternative models have been proposed either by modifying the right hand side of the Einstein field equations by considering the specific forms of energy -momentum tensor $T_{\mu\nu}$ or modifying the left hand side of Einstein field equations. The DE models which belong to the first category include quintessence [9], k-esssence [10], the family of chaplygin gas [11], holographic [12, 13], new agegraphic [14], etc. The models which belong to second category are the modified gravity that include $f(R)$ gravity [15], scalar -tensor theories [16]. The poineering study on scalar-tensor theories was done by Brans and Dicke [17] to incorporate Mach's principle into gravity which is known as Brans-Dicke (BD) theory. This was the first gravity theory in which the dynamics of gravity were described by a scalar field while spacetime dynamics were represented by the metric tensor. In BD theory, the gravitational constant G is replaced with the inverse of a time-dependent scalar field ϕ, namely, $\phi = (8\pi G)^{-1}$, and this scalar field couples to gravity with a coupling parameter ω. Since the BD theory passes the experimental tests from the solar system [18] and provides an explanation of the accelerated expansion of the Universe, it is worthwhile to discuss DE models in this framework.

In recent years, the holographic dark energy (HDE) [19–25] has been studied as a possible candidate for dark energy. This type of model was motivated by the holographic principle [12] which might lead to the quantum gravity to explain the events involving high energy scale. According to holographic principle, the number of degrees of freedom in a bounded system should be finite and has relations with the area of its boundary. For a system with size L and ultra-voilet (UV) cut-off Λ without decaying into a black hole, it is required that the total energy in a region of size L should not exceed the mass of a black hole of the same size, thus $L^3\rho_\Lambda \leq LM_P^2$. The largest L allowed is the one saturating this inequality, thus the holographic DE density $\rho_\Lambda = 3c^2M_P^2L^{-2}$ where c is a numerical constant and M_P is the reduced Planck mass $M_P^2 = 1/8\pi G$. It means that there must be a duality between UV cut-off and (infra-red) IR cut-off. Therefore, the UV cut-off is related to the vacuum energy and IR cut-off is related to the large scale of the Universe. The large scale of the Universe can be taken as, for example Hubble horizon, particle horizon or event horizon [19, 20]. Nowadays, the HDE model is the most efficient DE candidate among all the existing DE candidates which is also capable to explain the acceleration of the Universe. It has been shown that the HDE model is favored by the latest observational data including the sample of Type Ia supernovae (SN Ia), the shift parameter of the cosmic microwave background (CMB), and the baryon acoustic oscillation (BAO) measurement [26].

Li [19] has discussed three choices for the length scale L which is supposed to provide an IR cutoff. The first choice is to identify L with the Hubble radius, H^{-1}. In the formalism of HDE, the Hubble horizon is a most natural choice for the IR cut-off, but Hsu [20] in GR and Xu et al. [27] in BD theory have shown that the Hubble horizon as an IR cutoff is not a suitable candidate to explain the recent accelerated expansion. The second option is the particle horizon radius but it also does not give an accelerated expansion. The third choice is the identification of L with the radius of the future event horizon which gives the desired result, namely a sufficiently negative equation of state to obtain an accelerated universe.

In this paper, we consider the HDE model with IR cut-off as future event horizon in BD theory. Since most of the authors [21, 25, 28] have discussed HDE model in BD theory with power-law form of the BD scalar field. In a recent paper, Kumar and Singh [29] have revisited these papers and have proposed a logarithmic form of BD scalar field to discuss new age-graphic dark energy model in BD theory. The authors have also claimed to resolve the cosmic coincidence problem. Actually, it has been noticed in papers [21, 25, 28] that the power-law form of BD scalar field gives the constant value of deceleration parameter whatever may be matter content. Therefore, power-law assumption can not describe the phase transition of the Universe.

In this work, we revisit the study of Ref. [25] and consider HDE model in BD theory with logarithmic form of BD scalar field as proposed in Ref. [29]. We obtain the time -dependent deceleration parameter and equation of state parameter which describe the phase transition of the evolution of the Universe. We further discuss a cosmological model where the pressureless dark matter and HDE do not conserved separately but interact with each other. We also discuss the cosmic coincidence problem which has not been discussed in previous work [21,25]. It is found that the HDE model successfully resolves the cosmic coincidence problem in BD theory with event horizon.

The paper is organized as follows: In Sec 2, we present the field equations of HDE model in BD theory. We study the non-interacting holographic dark energy with IR cut-off as future event horizon. We calculate the equation of state parameter and deceleration parameter to discuss the nature of the evolution of the Universe. Section 3 deals with the interacting HDE model in BD theory. Section 4 gives the conclusion of the work.

2 Holographic dark energy in Brans-Dicke Theory

The modified Einstein-Hilbert (EH) action for the BD theory is given by [17]

$$S = \int d^4x\sqrt{g}\left(-\varphi R + \frac{\omega}{\varphi}g^{\mu\nu}\partial_\mu\varphi\,\partial_\nu\varphi + \mathcal{L}_m\right), \tag{1}$$

where R denotes the Ricci scalar curvature, ω is a dimensionless coupling parameter between scalar field and gravity called BD parameter and $\mathcal{L}_m$ represents the matter Lagrangian density. On re-defining the scalar field φ as

$$\varphi = \frac{\phi^2}{8\omega} \tag{2}$$

The action (1) of BD theory in the canonical form is given by [21, 30, 31]

$$S = \int d^4x\sqrt{g}\left(-\frac{1}{8\omega}\phi^2 R + \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi + \mathcal{L}_m\right), \tag{3}$$

where ϕ is a time-dependent scalar field called BD scalar field. The non-minimal coupling term $\phi^2 R$ replaces with the EH term R/G in such a way that $G_{eff}^{-1} = 2\pi\phi^2/\omega$, where G_{eff} is the effective gravitational constant as long as the dynamical scalar field ϕ varies slowly.

We consider a homogeneous and isotropic Friedmann-Robertson-Walker (FRW) Universe, which is described by the line element

$$ds^2 = dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\Phi^2)\right], \tag{4}$$

where $a(t)$ is the cosmic scale factor of the Universe and k is the curvature parameter with $k = -1, 0, 1$ corresponding to open, flat and closed Universes, respectively. We assume that the Universe is filled with perfect fluid containing pressureless dark matter (DM) (excluding baryonic matter) and HDE.

Taking the variation of action (3) with respect to the metric tensor, $g_{\mu\nu}$ for the line element (4), one can obtain the field equations for the non-flat universe containing HDE and DM as

$$\frac{3}{4\omega}\left(H^2+\frac{k}{a^2}\right)-\frac{1}{2}\frac{\dot{\phi}^2}{\phi^2}+\frac{3H}{2\omega}\frac{\dot{\phi}}{\phi}=\frac{\rho_m+\rho_h}{\phi^2}, \tag{5}$$

$$\frac{1}{4\omega}\left(2\frac{\ddot{a}}{a}+H^2+\frac{k}{a^2}\right)+\frac{H}{\omega}\frac{\dot{\phi}}{\phi}+\frac{1}{2\omega}\frac{\ddot{\phi}}{\phi}+\frac{1}{2}\left(1+\frac{1}{\omega}\right)\frac{\dot{\phi}^2}{\phi^2}=-\frac{p_h}{\phi^2}, \tag{6}$$

where $H=\frac{\dot{a}}{a}$ is the Hubble parameter, ρ_m is the DM energy density, ρ_h is the HDE energy density and p_h is the pressure of HDE. Here, an over dot denotes the derivative with respect to the cosmic time t. The wave equation for the scalar field ϕ is given by

$$\ddot{\phi}+3H\dot{\phi}-\frac{3}{2\omega}\left(\frac{\ddot{a}}{a}+H^2+\frac{k}{a^2}\right)\phi=0. \tag{7}$$

In frame work of BD theory, the HDE density has the form

$$\rho_h=3c^2\varphi L^{-2}, \tag{8}$$

where $\varphi=M_P^2=(8\pi G_{eff})^{-1}$. In canonical form $G_{eff}^{-1}=\frac{2\pi}{\omega}\phi^2$, the HDE density (8) now becomes

$$\rho_h=\frac{3c^2\phi^2}{4\omega L^2}. \tag{9}$$

It is to be noted that if the IR cut-off is taken as the Hubble horizon then the energy density of HDE and the critical density match identically. This situation generally arises in inflation scenario where $L=H^{-1}$. There are various other choices of IR cut-off for the cosmological length scale L available in the literature, such as particle horizon, event horizon, Ricci length, Granda-Oliveros cut-off etc. However, particle horizon is not suitable to derive the acceleration. A suitable choice of future event horizon as an IR cut-off was suggested by Li et al [32]. The cosmological length L for the event horizon is defined as

$$L=ar(t), \tag{10}$$

where the function $r(t)$ can be obtained from the relation

$$\int_0^{r(t)}\frac{dr}{\sqrt{1-kr^2}}=\int_t^{\infty}\frac{dt}{a(t)}=\frac{R_E}{a(t)}. \tag{11}$$

where R_E is the event horizon, defined by

$$R_E=a\int_t^{\infty}\frac{dt}{a(t)}=a\int_a^{\infty}\frac{da}{Ha^2}. \tag{12}$$

The general solution of $r(t)$ from Eq. (11) for non-flat FRW model is given by

$$r(t)=\frac{1}{\sqrt{k}}\sin y, \tag{13}$$

where, $y=\frac{\sqrt{k}R_E}{a(t)}$.
The critical energy density ρ_{cr} and the energy density of the curvature ρ_k are, respectively, defined as

$$\rho_{cr}=\frac{3\phi^2H^2}{4\omega}, \tag{14}$$

$$\rho_k=\frac{3k\phi^2}{4\omega a^2}. \tag{15}$$

It is useful to express the equations and physical quantities in the terms of fractional energy densities to analyse the results in better way. The fractional energy densities in their usual form are given as

$$\Omega_m = \frac{\rho_m}{\rho_{cr}} = \frac{4\omega\rho_m}{3\phi^2 H^2}, \tag{16}$$

$$\Omega_k = \frac{\rho_k}{\rho_{cr}} = \frac{k}{H^2 a^2}, \tag{17}$$

$$\Omega_h = \frac{\rho_h}{\rho_{cr}} = \frac{c^2}{H^2 L^2}. \tag{18}$$

Equation (18) can also be written as

$$HL = \frac{c}{\sqrt{\Omega_h}} \tag{19}$$

Now, differentiating (10) with respect to the cosmic time t, and using (13) and (19), we obtain

$$\dot{L} = HL + ar\dot{(}t) = \frac{c}{\sqrt{\Omega_h}} - \cos y. \tag{20}$$

Let us first consider the case where HDE and DM do not interact. In this case, the conservation equations for HDE and DM are respectively given by

$$\dot{\rho}_h + 3(1 + w_h)\rho_h H = 0, \tag{21}$$

$$\dot{\rho}_m + 3\rho_m H = 0, \tag{22}$$

where $w_h = \frac{p_h}{\rho_h}$ is the equation of state (EoS) parameter of HDE.
In order to discuss the physical behaviors of HDE model, we consider the following well-motivated ansatz for BD scalar field ϕ [29]

$$\phi = \phi_0 ln(\alpha + \beta a), \tag{23}$$

where $\phi_0 > 0$, $\alpha > 1$ and $\beta > 0$ are constants.

This logarithmic form of ϕ fulfills the requirement of slow variation of G as ϕ evolves slowly. It is important to notice that BD cosmology becomes standard cosmology when $\beta \to 0$. As we know that the universe exhibits phase transition from the past decelerated phase to the current accelerated phase. A time-dependent deceleration parameter (DP) is required to describe this phase transition. Therefore, the choice (23) of the logarithmic form of BD scalar field, which always gives time-dependent DP, is physically acceptable. Kumar and Singh [29] have discussed interacting new agegraphic dark energy (NADE) model with the Hubble horizon as an IR cut-off in the frame work of BD theory by assuming the above form of BD scalar field.

In this paper, we discuss the HDE cosmological model with future event horizon as an IR cut-off in the frame work of BD theory as Li et al [32] suggested that this IR cut-off is suitable to describe the accelerated expansion of the universe.

On taking the derivative of (23) with respect to time, we get

$$\dot{\phi} = \phi_0 \frac{\beta a H}{(\alpha + \beta a)}, \tag{24}$$

$$\ddot{\phi} = \phi_0 \left\{ \frac{\beta a \dot{H}}{(\alpha + \beta a)} + \frac{\beta a H^2}{(\alpha + \beta a)} - \frac{\beta^2 a^2 H^2}{(\alpha + \beta a)^2} \right\}. \tag{25}$$

By use of (20) and (24), Eq.(9) gives

$$\dot{\rho}_h = 2H\rho_h \left(-1 + \frac{\sqrt{\Omega_h}}{c} \cos y + \frac{\beta a}{(\alpha + \beta a)\ ln(\alpha + \beta a)} \right). \tag{26}$$

Using (26) into (21), we obtain the EoS parameter for HDE

$$w_h = -\frac{1}{3} - \frac{2\beta a}{3(\alpha+\beta a)\, ln(\alpha+\beta a)} - \frac{2\sqrt{\Omega_h}}{3c}\cos y. \tag{27}$$

The EoS w_h is bounded from below by

$$w_h = -\frac{1}{3} - \frac{2\beta a}{3(\alpha+\beta a)\, ln(\alpha+\beta a)} - \frac{2\sqrt{\Omega_h}}{3c}. \tag{28}$$

From (27), we observe that for $\beta = 0$ the EoS parameter w_h of HDE reduces to its respective form of non-flat standard cosmology [33], which is given by

$$w_h = -\frac{1}{3} - \frac{2\sqrt{\Omega_h}}{3c}\cos y. \tag{29}$$

It is to be noted that for power-law form $\phi \propto a^m$ of BD scalar field [21, 25], the second term in the value of w_h of Eq.(28) is a constant term $\frac{2m}{3}$ whereas we get a time-dependent term $\frac{2\beta a}{3(\alpha+\beta a)\, ln(\alpha+\beta a)}$ due to logarithmic form of BD scalar field. Therefore, the value of w_h in our model is more dynamic in comparison to power-law in BD theory.

Let us discuss the behavior of EoS parameter w_h of HDE as obtained in Eq.(28). It is clear that w_h always has a negative value such that $w_h < -\frac{1}{3}$. The value of w_h at the beginning of the evolution, i.e., $a = 0$ is same as in Eq. (29) for GR because the second term is zero at $a = 0$. Thus, we find a negative value of w_h in the very early Universe. It is observed that the second term attains to maximum value during the evolution and approaches to zero in late time evolution. The maximum value depends on sufficiently small values of α and large values of β during the process of evolution. However, the maximum value only depends on the parameter α and it is found that $\max\left\{\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}\right\} \to 1$ as $\alpha \to 1$, i.e., $w_h < -1$. Therefore, w_h may cross phantom divide line for this condition. It may also possible that w_h crosses the phantom divide line for $\sqrt{\Omega_h} > c[1 - \frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}]$, i.e., for both conditions HDE model may cross the phantom divide line ($w_h = -1$) and approaches to the phantom region.

It is also interesting to note that as the logarithmic term converges to zero in the late-time of evolution and also $\cos y \to 1$ in late-time as $a \to \infty$, the EoS parameter starts behaving like its respective form in standard general relativity [33] and it will depend only on the values of Ω_h and c. The form of w_h in the late-time is given by

$$w_h = -\frac{1}{3} - \frac{2\sqrt{\Omega_h}}{3c}. \tag{30}$$

In this model, the value of parameter c determines the property of HDE in late time. Since, the observation predicts $\Omega_h \to 1$ for the present time, therefore, at $c = 1$, w_h approaches to -1, i.e., our model behaves like cosmological constant. We get $w_h > -1$ but less than $-1/3$ at $c > 1$, i.e., our model shows the quintessence region and if $c < 1$, we get $w_h < -1$, i.e., the phantom type behaviour occur. Thus, we conclude that when $c > 1$, $c = 1$ and $c < 1$, one can generate quintessence, cosmological constant and phantom respectively for non-interacting HDE model in BD theory.

Now, we study the behaviour of deceleration parameter (DP) to discuss the evolution of the Universe. The DP, $q = -\frac{a\ddot{a}}{\dot{a}^2} = -1 - \frac{\dot{H}}{H^2}$, can be obtained after dividing Eq. (6) by H^2, and using Eqs. (9), (17), (19), (24) and (25), which is given as

$$q = \frac{1 + \Omega_k + 3w_h\Omega_h + \frac{4\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)} - \frac{2\beta^2 a^2}{(\alpha+\beta a)^2\, ln(\alpha+\beta a)} + \frac{2(\omega+1)\beta^2 a^2}{(\alpha+\beta a)^2\, [ln(\alpha+\beta a)]^2}}{2\left(1 + \frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}\right)}. \tag{31}$$

The term $\frac{\beta^2 a^2}{(\alpha+\beta a)^2\, ln(\alpha+\beta a)}$ has the same behaviour as the term $\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}$ except it has maximum value lies in (0, 0.41) depending on the value of α. The BD parameter ω also plays

an important role in the value of q. The solar system experiment Cassini gave a very high bound on ω as $|\omega| > 40000$ [34,35], whereas the cosmological observations provide the relatively lower bounds on ω [36–39]. The observations suggest that ω has the large value so the last term of numerator of Eq. (31) containing ω will dominate during the evolution of the Universe. This shows q may attain some positive value, i.e., the decelerated expansion of the Universe may occur during the evolution. Thus, HDE model explains the matter dominated phase of the Universe. As in the late-time of evolution, $w_h \to -1$ and the terms $\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}$ and $\frac{\beta^2 a^2}{(\alpha+\beta a)^2\, ln(\alpha+\beta a)}$ converge to zero as $a \to \infty$. Then, the value of q in the late time of evolution is obtained as

$$q \approx \frac{1+\Omega_k-3\Omega_h}{2}. \tag{32}$$

Since the observations recommend that our present Universe is almost flat, i.e., $k = 0$ ($\Omega_k = 0$), Eq.(32) gives $q \approx \frac{1-3\Omega_h}{2}$. Thus, it is observed that q is negative for $\Omega_h > 1/3$, i.e., the accelerated expansion is obtained for $\Omega_h > 1/3$. It can also be noticed that if we consider the open Universe, i.e., $k < 0$ ($\Omega_k < 0$), the accelerated expansion can be obtained more easily. Even for the closed geometry case of the Universe we can also get an accelerated expansion of the Universe but for this we must have a very large value of Ω_h which will give a negative value of q. Thus, we can conclude that the HDE model describes the phase transition from early time inflation to the matter dominated phase and then matter dominated phase to late time accelerated phase.

Let us consider the cosmological coincidence problem which was raised first time by Steinhardt [40, 41]. The problem may be resolved by making the density ratio $r_1 = \frac{\rho_m}{\rho_h}$ is of order unity, i.e., $(r_1)_0 \sim \mathcal{O}(1)$ for a wide range of initial condition. The second way is that either r_1 converses to a constant value or evolve very slowly in late-time of evolution. From (5), the energy density ratio is given by

$$r_1 = -1 + \frac{1}{\Omega_h}\left[\Omega_k + 1 - \frac{2\,\omega\beta^2 a^2}{3(\alpha+\beta a)^2\,[ln(\alpha+\beta a)]^2} + \frac{2\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}\right]. \tag{33}$$

Therefore, we obtain a time-dependent value of r_1. At the beginning of evolution the value of r_1 is $\left\{-1+\frac{(\Omega_k+1)}{\Omega_h}\right\}$ as the last two terms vanish at $a = 0$. In the late time of evolution we obtain the same expression of r_1 as the last two terms approaches to zero as $a \to \infty$. Now, the evolution of energy density ratio r_1 can be obtained as

$$\dot{r}_1 = 3r_1 H w_h. \tag{34}$$

According to ΛCDM model, r_1 evolves as $|\frac{\dot{r}_1}{r_1}|_0 = 3H_0$. Throughout the paper the subscript zero represents the present value of the quantity. Since, for $c = 1$ our model shows $w_{h0} = -1$ in late-time, therefore, we get $|\frac{\dot{r}_1}{r_1}|_0 = 3H_0$ which is same as for ΛCDM model. This shows that there is no reduction in the acuteness of the coincidence problem. Since, the EoS parameter w_h is time-dependent value, therefore, the less acute coincidence problem can be obtained if we have a quintessence like EoS parameter ($w_h > -1$). Also, we may achieved $w_{h0} > -1$ for $\sqrt{\Omega_h} < \frac{c}{\cos y}[1 - \frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}]$. Since, the second term converges to zero as $a \to \infty$, we can get quintessence like EoS parameter more conveniently due to the logarithmic form as compare to power law form of BD scalar field whereas we get the constant second term. Thus, we can conclude that this logarithmic form of BD scalar field is more appropriate to achieve a less acute coincidence problem. Now, let us assume $w_{h0} = -2/3$ we obtained $|\frac{\dot{r}_1}{r_1}|_0 = 2H_0$. Clearly, it shows less acuteness in the coincidence problem as compare to the ΛCDM model. But, the problem is more acute in the case of phantom like EoS parameter ($w_{h0} < -1$). This case shows the more complex condition of coincidence problem as compare to the ΛCDM model.

3 Interacting HDE in Brans-Dicke Cosmology

In this section, we extend our study to the case where both dark components, the pressureless DM and the HDE, interact with each other. If we proceed to consider a scenario of interacting dark energy, ρ_m and ρ_h do not satisfy independent conservation laws, they instead satisfy

$$\dot{\rho}_h + 3H(1+w_h)\rho_h = -\Gamma, \tag{35}$$

and

$$\dot{\rho}_m + 3H\rho_m = \Gamma, \tag{36}$$

where $\Gamma = 3b^2H(\rho_m+\rho_h)$ is a particular interacting term [42] with the coupling constant b^2. This interacting term can be rewritten in terms of ratio of density parameter $r_1 = \rho_m/\rho_h$ as

$$\Gamma = 3b^2H\rho_h(1+r_1), \tag{37}$$

Using (33) and (37) into (35), the EoS parameter of HDE is given by

$$\begin{aligned} w_h &= -\frac{1}{3} - \frac{2\sqrt{\Omega_h}}{3c}\cos y - \frac{2\beta a}{3(\alpha+\beta a)\, ln(\alpha+\beta a)} \\ &- \frac{b^2}{\Omega_h}\left[1+\Omega_k + \frac{2\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)} - \frac{2\omega\beta^2a^2}{3(\alpha+\beta a)^2\, [ln(\alpha+\beta a)]^2}\right], \end{aligned} \tag{38}$$

which is a time-dependent value. It is to be noted that the term $\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}$ has the same behaviour as discussed earlier for non-interacting case. It is easy to find out that, in the limit of $\beta \to 0$, the standard cosmology is recovered. In the beginning of the evolution, the term $\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}$ is zero and hence the EoS parameter of HDE (38) gives

$$w_h = -\frac{1}{3} - \frac{2\sqrt{\Omega_h}}{3c}\cos y - \frac{b^2(1+\Omega_k)}{\Omega_h}, \tag{39}$$

which is same as the standard non-flat HDE model in BD theory. We find that w_h is always negative and less than $-1/3$ in the early of the evolution which shows the inflation in early time. The solar system experiments [34] predict a very high bound value of ω which is $|\omega| > 40000$. However, Acquaviva and Verde [43] found that ω may be smaller than 40000 in cosmological scale. Due to a large value ω suggested by the experiments, the last term containing ω will dominate during early phase of the evolution of the universe. Thus, we observe a positive value of ω. Thus, the decelerated phase occurs during the evolution of the universe. It means that the universe passes through the matter-dominated phase. During late time evolution, EoS parameter of HDE becomes

$$w_h = -\frac{1}{3} - \frac{2\sqrt{\Omega_h}}{3c} - \frac{b^2(1+\Omega_k)}{\Omega_h}, \tag{40}$$

which gives a negative value. Analysing w_h in (40), one can observe that w_h will definitely cross the phantom divide line in the late time evolution. The late time value of w_h depends on the values of coupling constant b^2, c, Ω_h and Ω_k.

Now, dividing the Eq. (6) by H^2 and substituting the value of w_h from Eq. (40), we get the following value of deceleration parameter for interacting HDE.

$$\begin{aligned} q &= \frac{(1+3b^2)(1+\Omega_k) - \Omega_h\left(1+\frac{2\sqrt{\Omega_h}}{c}\cos y\right)}{2\left(1+\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}\right)} \\ &+ \frac{\frac{2\beta a(-\Omega_h+3b^2+2)}{(\alpha+\beta a)\, ln(\alpha+\beta a)} - \frac{2\beta^2a^2}{(\alpha+\beta a)^2\, ln(\alpha+\beta a)} + \frac{2(\omega+1-3\omega b^2)\beta^2a^2}{(\alpha+\beta a)^2\, [ln(\alpha+\beta a)]^2}}{2\left(1+\frac{\beta a}{(\alpha+\beta a)\, ln(\alpha+\beta a)}\right)}, \end{aligned} \tag{41}$$

From (41), we observe that the last term will dominate during the evolution due to the large value of the BD parameter ω provided $1+3b^2>0$, which means that q becomes positive and hence it describes the decelerated phase. In the late-time of the evolution the terms $\frac{\beta a}{(\alpha+\beta a)\,ln(\alpha+\beta a)}$, $\frac{\beta^2 a^2}{(\alpha+\beta a)^2\,ln(\alpha+\beta a)}$ and $\frac{\beta^2 a^2}{(\alpha+\beta a)^2\,[ln(\alpha+\beta a)]^2}$ converge to zero and $\cos y$ converge to 1, the value of q is given by

$$q=\frac{(1+3b^2)(1+\Omega_k)-\Omega_h(1+\frac{2\sqrt{\Omega_h}}{c})}{2}. \tag{42}$$

From above we observe that q is negative for $b^2<\frac{\Omega_h(1+\frac{2\sqrt{\Omega_h}}{c})}{3(1+\Omega_k)}-\frac{1}{3}$.

Let us discuss coincidence problem in interaction HDE model. Using (36), (37) and (38) the evolution of r_1 can be expressed as

$$\dot{r}_1=3Hr_1\left[w_h+\frac{b^2(1+r_1)^2}{r_1}\right]. \tag{43}$$

In the above expression the value within the bracket can be positive or negative but for a suitable value of b^2, we can get $|w_h+\frac{b^2(1+r_1)^2}{r_1}|<<|w_h|$. Thus, from Eqs.(34) and (43) we can conclude that the energy density ratio r_1 may evolve more slowly in interacting HDE model as compare to non-interacting HDE model. This imply that the interaction between DM and HDE plays a vital role to discuss the coincidence problem. As we know that for getting the soft coincidence, the model must satisfy the condition $|\frac{\dot{r}_1}{r_1}|_0\leq H_0$. In our model we can achieve the soft coincidence if b^2 satisfy the condition $b^2\leq\frac{(1-3w_{h0})r_{10}}{3(1+r_{10})^2}$. According to the present observational values $r_{10}=3/7$ and $w_{h0}=-1$, we get $b^2\leq\frac{7}{25}$. Thus soft coincidence can be achieved at present if $b^2\leq\frac{7}{25}$. It can also be concluded that the smaller the value of b^2, the energy density ratio may evolve more slowly. This explanation can resolve the problem of cosmic coincidence and it can be checked by taking any suitable small value of b^2 along with $w_{h0}=-1$. This represents that the variation in r_1 is more slow as compare to the conventional ΛCDM model. Thus, the coupling constant b^2 plays an important role to resolve the cosmic coincidence problem and also this small value of coupling constant is compatible with the observations. This is also analyzed by Feng et al. [44]. Thus, we observe that interacting HDE along with the logarithmic form of BD scalar field in the framework of BD theory may capable to resolve the cosmic coincidence problem.

4 Conclusion

In this paper, we have studied non-interacting and interacting HDE model with future event horizon as an infra-red cut-off in the framework of BD theory. Motivated by work [29] we have considered the logarithmic form of BD scalar field to discuss the early and late time behaviour of the Universe. This form of BD scalar field gives a time-varying deceleration parameter for non-interacting and interacting HDE models irrespective of the matter content. We have discussed the dynamical view of early and late-time evolution of the Universe with the help of EoS parameter and deceleration parameter. We have also discussed the cosmic coincident problem. The result of both, non-interacting and interacting models are summarize below.

In the first case where we have considered the non-interacting HDE model, we have observed that the EoS starts behaving like as its respective form in GR. In late-time of evolution at $c=1$ our model behaves like cosmological constant, at $c>1$ it shows the quintessence region and it mimic like phantom type at $c<1$. Initially the value of DP is negative but during the evolution the very high bound on BD parameter ω dominates the other terms and may attain the positive value and it is also observed that in late-time it again attain a negative value, i.e.,

it shows the phase transition from decelerated to accelerated Universe during the evolution. We have also discussed the cosmic coincidence problem. In this by assuming $\omega_h = -2/3$, i.e., $|\frac{\dot{r}_1}{r_1}| = 2H_0$, we get the less acute coincidence problem as compare to ΛCDM.

In the second case where we consider the interacting HDE model, we get a time-dependent value of EoS parameter. In early-time and late-time it behaves same as its respective form of GR. In beginning it gives a negative value which is $< -1/3$, this shows the early-time inflation. Due to the presence of the BD parameter ω, which have a very high positive value(according to the observations), in the EoS parameter the value of ω_h may get some positive value at any time during the evolution depending upon the values of α, β and ω. This means that during the evolution of the Universe the decelerated phase of the Universe may occur. In the late-time of the evolution it again gives the negative value which is $< -1/3$. This shows the acceleration of the Universe at late-time. Also, in this case we have obtained a time-depending value of the DP. The presence of the BD parameter shows that during the evolution of the Universe under the condition $1 + 3b^2 > 0$, q may attain a positive value. In late-time q will show the accelerated Universe under the condition $b^2 < \frac{\Omega_h(1+\frac{2\sqrt{\Omega_h}}{c})}{3(1+\Omega_k)} - \frac{1}{3}$. Thus, we conclude that the interacting HDE case in our model shows the phase transition from deceleration to acceleration, which is a good harmony with the current observations. Also in this case we discuss the cosmic coincidence problem. Since the coupling constant plays a vital role to discuss the cosmic coincidence problem and the smaller value of b^2 shows that the energy density ratio may evolve more slowly. Here we observe that the soft coincidence can be achieved if $b^2 \leq \frac{7}{25}$. Thus we can conclude that in the framework of BD theory the interacting HDE with the logarithmic form of BD scalar field may capable to resolve the cosmic coincidence problem.

References

[1] A. G. Riess et al., *Astron. J.* **116**, 1009 (1998).

[2] S. Perlmutter, et al., *Astrophys. J.* **517**, 565 (1999).

[3] C. L. Bennett et al., *Astrophys. J.* **148**, 1 (2003).

[4] E. Komatsu et al., *Astrophys. J. Suppl.* **192**, 18 (2011).

[5] W. J. Percival et al., *Mon. Not. R. Astro. Soc.* **401**, 2148 (2010).

[6] S. M. Carroll, *Living Rev. Relativity* **4**, 1 (2001).

[7] S. Weinberg, *Reviews of Modern Physics* **61**, 1 (1989).

[8] E.J. Copeland, M. Sami, S. Tsujikawa *Int. J. Mod. Phys. D*, **15**, 1753 (2006).

[9] S. M. Carroll, *Phys. Rev. Lett.* **81**, 3067 (1998).

[10] T. Chiba, T. Okabe, M. Yamaguchi, *Phys. Rev.D* **62**, 023511 (2000).

[11] A. Y. Kamenshchik et al., *Phys. Lett. B* **511**, 265 (2001).

[12] G. 't Hooft, arXiv:gr-qc/9310026 (1993).

[13] L. Susskind, *J. Math. Phys.* **36**, 6377 (1995).

[14] H. Wei and R. G. Cai, *Phys. Lett. B* **660**, 113 (2008).

[15] S. Copozzielo, *Int. J. Mod. Phys. D* **11**, 483 (2002).

[16] L. Amendola, *Phys. Rev. D* **60**, 043501 (1999).

[17] C. Brans and R. H. Dicke, *Phys. Rev.* **124**, 925 (1961).

[18] B. Bertotti, L. Iess and P. Tortora, *Nature* **425**, 374 (2003).

[19] M. Li, *Phys. Lett. B* **603**, 1 (2004).

[20] S. D. Hsu, *Phys. Lett. B* **594**, 1 (2004).

[21] M.R. Setare, *Phy. Lett. B* **644**, 99 (2007).

[22] M.R. Setare, *Eur. Phys. J. C* **50**, 991 (2007).

[23] N. Cruz, S. Lepe, F. Pena, J. Saavedra, *Phys. Lett. B* **669**, 271 (2007).

[24] J. Zhang, X. Zhang, H. Liu, *Phys. Lett. B* **659**, 26 (2008).

[25] A. Sheykhi, *Phys. Lett. B* **681**, 205 (2009).

[26] Q. Wu, Y. G. Gong, A. Z. Wang, and J. S. Alcaniz, *Phys. Lett. B* **659**, 34 (2008).

[27] L. Xu, W. Li and J. Lu, *Eur. Phys. J. C* **60**, 135 (2009).

[28] N. Banerjee and D. Pavón, *Phys. Lett. B* **647**, 477 (2007).

[29] P. Kumar and C. P. Singh, *Astrophys. Space Sci.* **362**, 52 (2017).

[30] M. Arik and M. C. Calik, *Mod. Phys. Lett. A* **21**, 1241 (2006).

[31] M. Arik, M. C. Calik and M. B. Sheftel, *Int. J. Mod. Phys. D* **17**, 225 (2008)

[32] M. Li, X. D. Li, S. Wang and X. Zhang, *J. Cosmol. Astrophys. Phys.* **906**, 036 (2009).

[33] Q. G. Huang and M. Li, *J. Cosmol. Astrophys. Phys.* **0408**, 013 (2004).

[34] B. Bertotti, L. Iess and P. Tortora, *Nature* **426**, 374 (2003).

[35] C. Will, *Living Rev. Relativity* **9**, 3 (2006).

[36] L. Xu et al., *Mod. Phys. Lett. A* **25**, 1441 (2010).

[37] V. Acquaviva et al., *Phys. Rev. D* **71**, 104025 (2005).

[38] F.-Q. Wu and X. Chen, *Phys. Rev. D* **82**, 083003 (2010).

[39] Y.-C. Li, F.-Q. Wu and X. Chen, *Phys. Rev. D* **88**, 084053 (2013).

[40] P. Steinhardt, in *Critical Problems in Physics*, edited by V. L. Fitch and D. R. Marlow (Princeton University Press, Princeton, NJ, 1997).

[41] I. Zlatev, L. Wang and P. J. Steinhardt, *Phys. Rev Lett.* **87**, 5 (1999).

[42] K. Karwan, JCAP bf05, 011 (2008).

[43] V. Acquaviva and L. Varde, JCAP **12**, 001 (2007).

[44] C. Feng et al., *Phys. Lett B* **665**, 111 (2008).

Int J Syst Assur Eng Manag
DOI 10.1007/s13198-017-0633-5

ORIGINAL ARTICLE

Integrated scheduling of part and tool in a flexible manufacturing system using modified genetic algorithm

Naveen Kumar[1] · Pankaj Chandna[2] · Dheeraj Joshi[3]

Received: 30 January 2017 / Revised: 3 April 2017
© The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2017

Abstract Scheduling problems in an FMS have been considered as complex optimization problems whose solution by conventional techniques requires a great deal of efforts and time. In this paper, a simultaneous loading and scheduling of part and tool has been proposed for a flexible manufacturing system which has identical machines and a common tool magazine. All the tools are stored in the common tool magazine, and shared among the different machines through a material handling system. Each tool type is single in number. A modified genetic algorithm (MGA) with three parent crossover and a mutation operator is used to find the optimal solution of the loading and scheduling problem. The MGA uses an algorithm which is based on Giffler and Thompson procedure with a heuristic approach to resolve the job conflict and generate an active feasible schedule. The performance of the proposed algorithm is analyzed by comparing the makespan results with the results existing in literature. It is observed that the MGA yields better results than the algorithms reported so far. Furthermore, efficiency of MGA improves as the problem size increases.

Keywords Loading · Scheduling · Flexible manufacturing system · Modified genetic algorithm

Abbreviations

FMS	Flexible manufacturing system
AGVs	Automated guided vehicles
ASRS	Automated storage and retrieval system
PDRA	Priority dispatching rules algorithm
CTM	Central tool magazine
GA	Genetic algorithm
GADG	Genetic algorithm with dominant genes
ACO	Ant colony optimization
PNs	Petri nets
ASMEA	Symbiotic evolutionary asymmetric multileveled algorithm
WIP	Work in process
SAA	Simulated annealing algorithm
FMC	Flexible manufacturing cell

List of symbols

est_{ik}	Earliest start time of kth operation of ith job
eft_{ik}	Earliest finishing time *k*th operation of *i*th job
DT	Datum time
N	Number of jobs
K	Number of operations
t_{ik}	Processing time of *k*th operation of *i*th job
IP	Initial population size
s	Population size of the selection pool A
J_i	Job number ($i = 1$ to n)
J_{ik}	*k*th operation of *i*th job
COJ	Conflict of jobs
M_j	Machine number ($j = 1$ to m)
m	Number of machines
MT	Makespan time
p	Mutation probability
MGA	Modified genetic algorithm
$MAXGEN$	Maximum number of generations for MGA
$P.O.J$	Processed operation of job
JA	Job assigned

✉ Naveen Kumar
agrawal.naveen1@gmail.com

[1] Sir Padampat Singhania University, Udaipur, India

[2] NIT Kurukshetra, Kurukshetra, India

[3] Delhi Technological University, Delhi, India

Published online: 27 May 2017

1 Introduction

To meet the challenge of fast changing demands, the manufacturer must have to produce a large number of product types with minimum lead time. The advent of computer numerical controlled machines has resulted in the development of flexible manufacturing systems (FMS). A Flexible manufacturing system facilitates the production of variety of part types in smaller batch size with minimum setup time. Such system consists of number of general purpose numerically controlled machines, interconnected by automated material handling system and the control is being done by a central computer. Using FMS for production, flexibility of job shops and efficiency of mass production can be enhanced; however, setting up an FMS requires a lot of investments, hence, the efficient solutions to the various decision problems are key to avoid its underutilization (Turkcan et al. 2007).

There are four stages of decision problems for successful installation and implementation of an FMS: designing, planning, scheduling and control (Stecke 1983). Loading and scheduling problem is to allocate parts (or operations) and required tools to machines with exact time span (Roh and Kim 1997). To fully exploit the benefits of an FMS, careful attention must be paid to operational issues like loading, scheduling and control (Gnanavel et al. 2010). Scheduling of part and tool without considering other resources in the system, e.g. material-handling system like automated guided vehicles (AGVs) and automated storage and retrieval system (ASRS), lowers the efficiency and flexibility of FMS production activity. Many FMSs employ automated guided vehicle in material-handling system in order to improve the flexibility and efficiency of its production activity. The operation and control of AGVs is done with the help of well-designed vehicle management system.

FMS loading and scheduling problem has drawn considerable attention of researchers both from academics and industries in past four decades. FMS scheduling is much more complicated than job shop scheduling due to inherent flexibility of FMSs. Literature is replete with variety of scheduling problems concerning machining environment, job description, and objective function (French 1982; Brucker 1995). Part and tool flows in a FMS are two dynamic entities and their management is important for its efficient operation (Prabaharan et al. 2006). Scheduling of parts without considering the tool flow may lead to an inefficient working of FMSs, thereby preventing an FMS reaching to its fullest potential and make it 'inflexible' in practice (Gray et al. 1993; Veeramani et al. 1992; Selim and Ozkan 2001). In most of the existing works on loading and scheduling problems, the part and tool flow are considered as separate issues owing to computation complexity, and often, the effect of one of the pair on the other is neglected (Prabaharan et al. 2006). This might lead to suboptimal solution to the loading and scheduling problem, as solution to one issue affect the solution of other.

Scheduling problem in an FMS is generally solved using any of the two approaches: part movement approach, or tool movement approach (Roh and Kim 1997). In part movement approach, the tools are loaded on the machines and part moves to different machines depending on the availability of the required cutting tool on the machine. In tool movement policy, all the operations of a part are performed on the same machine and the required cutting tools are moved to the machine. The first approach is generally used when the cost of the cutting tools is considered to be insignificant in comparison to the cost of the parts. Most of the research work related to scheduling in an FMS has used the part movement approach (Mukhopadhyay and Nandi 1999). But when the tool cost contributes significantly to the cost of the part; sometimes as high as 25–30% (Selim and Siraceddin 1999); the second approach is being used. Due to limited tool budget, it is essential to find out the tool copy configuration for the best system performance (Jun et al. 1999). Therefore, by adopting appropriate tooling strategies and economizing on the tooling cost, large reduction on the tooling cost are feasible (Gray et al. 1993). In tool flow approach, automated transport of tools to the machines has been considered. The added cost of tool automation improves the economic efficiency of an FMS (Gray et al. 1993). Although the scheduling problems have been studied under the part movement policy extensively in many papers, there are relatively few articles on the problems under the tool movement policy (Roh and Kim 1997). A few researchers in their surveys on the tool management issues of automated manufacturing systems established that the lack of tooling considerations has resulted in the poor performance of these systems and stressed for tool scheduling (Gray et al. 1993; Veeramani et al. 1992). Most of the existing work on loading and scheduling in an FMS solve the problem sequentially, i.e. first the loading problem is solved, and based on the results obtained from the solution of loading problem, scheduling problems are solved. This approach does not give the optimal solution to the loading and scheduling problem, as the solution of loading problem becomes a limitation for the scheduling problem (Kim et al. 2007). Considering these aspects, in this paper the part scheduling and tool allocation problem has been solved simultaneously. The concept of common tool magazine (CTM) that shares with and serves for several machines, reduce the cost of duplicating tools in each and every machining centre, is of particular interest in each FMS.

Scheduling of parts with identical machines in an FMS even without additional resources has been considered to be NP-hard and requires heuristic approach for solution (Agnetis et al. 1997). An FMS scheduling problem is generally solved using optimization and heuristic based methods, however, Optimization methods, e.g. branch and bound, integer programming, dynamic programming etc., become cumbersome to solve as the problem size increases. Therefore, efficient heuristic methods need be developed for large sized problems. Priority dispatching rule algorithms (PDRA) are the simple heuristics which can be easily applied to any FMS scheduling problem. Giffler and Thompson (GT) algorithm (Giffler and Thompson 1960) developed for the job shop manufacturing environment can also provide quality results for an FMS scheduling problem (Nasciment 1993). Prabaharan et al. (2006) used combined PDRA and simulated annealing algorithm for the scheduling and sequencing of parts and tools in a flexible manufacturing cell (FMC). Udhayakumar and Kumanan (2010) used Ant colony optimization (ACO) algorithm to schedule the tools and parts with the objective of minimizing the makespan in an FMS having identical machines. Fathi and Barnette (2002) solved the scheduling problem of part and tool in an FMS with identical machine environment in order to minimize makespan using three heuristic methods.

Artificial Intelligence techniques have drawn considerable attention of researchers for solving scheduling problem in FMS. One of such intelligent probabilistic search technique is genetic algorithm (GA) (Holland 1975). According to Goldberg (1989), GAs can be applied to treat the complexity levels required to provide adaptive search at the requisite robustness. Genetic algorithm has been successfully used to solve verities of optimization problems including problems related to manufacturing, i.e. scheduling, process planning and system design. Ponnambalam et al. (2001) stated that GA is the most popular type of evolutionary algorithm to solve FMS scheduling problems. Keung et al. (2003) in their work proposed a GA to solve the sequencing problem in an FMS having one material handling device, with an objective to minimize the penalty cost. Shankar et al. (2005) designed multi-objective evolutionary algorithm equipped with a mechanism to generate parallel diverse optimal solutions, for scheduling of an FMS. Chan et al. (2008) used Genetic Algorithm with Dominant Genes (GADG) to solve the scheduling problem in an FMS. The results obtained for the minimization of makespan, were compared with results obtained by other techniques like Ant Colony Optimization and Petri Nets (PNs). The results obtained by the proposed algorithm were found to be better. Kim et al. (2007) considered an FMS with four types of flexibility: machine, tools, processing and sequencing. They used Symbiotic Evolutionary Asymmetric Multileveled Algorithm (ASMEA) to solve the scheduling problem in the said FMS environment. They reported that the results obtained were of high quality and the speed of convergence of the algorithm was also fast. Hsu et al. (2008) solved sequencing problem of cyclic tasks in an FMS by using GAs and PNs. They selected work in process (WIP) as performance parameter. They concluded that the results obtained by GA in 75% of the cases were equivalent to that of obtained by best heuristics available in the literature. Gang and Wu (2004) used hybrid approach (GA and PN) to solve sequencing problems in FMS. Reddy and Rao (2006) used GA and a heuristic for simultaneous scheduling of machines and AGVs (automated guided vehicles). Balin (2011) used GA to solve scheduling problem with non-identical parallel machines. The author used new crossover operator and optimality criteria for minimization of makespan. Godinho et al. (2014) in their work presented a comprehensive review of the available literature on GA applied to FMS scheduling. The analysis of the literature is done based on the proposed six classifications. Kaplanoglu (2016) in his work proposed an object oriented approach along with simulated annealing optimization algorithm for multi-objective flexible job-shop scheduling problem. Wu et al. (2017) in his work proposed genetic algorithm with a new chromosome representation scheme to solve the distributed flexible job-shop scheduling problem.

In GA evolutionary process; crossover and mutation are the important search operators. Most of the work reported on application of GAs in FMS scheduling problem, have considered two parent crossover method. Elsayed et al. (2014) proposed a new GA with multi-parent crossover with a diversity operator. They tested this algorithm on number of optimization problems and the results show its better performance.

This work proposes an FMS environment consisting of 'm' identical machines with a CTM. In case, any two or more machines require same tool at the same time, one of the machine is served with the tool in order to resolve the job conflict, while the other machines have to wait, till the tool completes the job on the assigned machine. The makespan has been used as performance measure in this study, because, it represents a long term or steady state performance measure for many manufacturing systems. In this work, we have used the GA having three-parent crossover with a mutation operator.

The paper is structured as follows: the formulated problem is described in Sect. 2. In Sect. 3 a new algorithm Active Schedule (ACT_SCH) and MGA are discussed. Discussion of results and sensitivity analysis is presented in Sect. 4. Finally, the conclusions are given in Sect. 5.

2 Problem structure

In many manufacturing systems tools are stored at a common place called common tool magazine. These tools are shared with machines in the system, through a material handling system in order to reduce the tool inventory. In this paper, problem of simultaneous scheduling of tools and part has been attempted for a flexible manufacturing system with a common tool magazine with an objective to minimize the makespan. Description of the problem environment, assumptions and performance measure are discussed in the section below.

2.1 FMS environment

The FMS in the problem has identical machines. The machines are capable of processing a group of parts and part verity. None of machines has dedicated tool magazine. All the tools are stored in the CTM and shared among all the machines through single tool grip. There is a limitation on the quantity of each tool type in CTM and it is kept as one in this problem. The FMS environment is shown in Fig. 1.

2.2 Assumptions

- Machines are capable of doing all operations of any of the jobs.
- Each job has a number of operations and individual operation requires a specific tool.
- Sequence of operations and respective tools vary from job to job.
- The operations sequence and type of tool required for each operation with processing time of operations are pre-specified.
- One job can be processed on one machine/tool at a time.
- All the operations of a job are processed during a single machine visit.

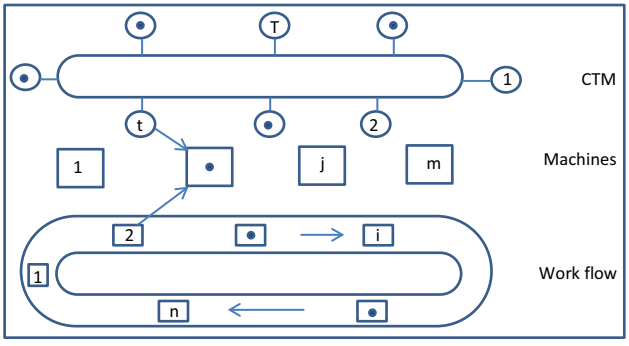


Fig. 1 Working environment of FMS

- Operation cannot be interrupted, i.e., each operation once started must be completed.
- A job does not visit the same machine twice.
- Availability of tools in each verity is considered as one and can be used more than once for any job.
- The operation time of a job includes the loading, unloading, tool changeover and setup times (both tool and job) along with the processing time.
- Once the operation is over the tool returns to the common tool magazine with negligible transfer time and is available for the next operation.

2.3 Performance measure

Considering the high initial cost of investment in FMSs, it is important for the manufacturer to maximize the machine utilization. In this context, minimum makespan that increases the utilization of machines is considered as the appropriate measure of system performance.

The formulae used for makespan are given as below:

Job completion time

$$C_i = \sum_{k=1}^{K} t_{ik} \tag{1}$$

$$\text{Makespan} = \max(C_1, C_2, \ldots, C_n) \tag{2}$$

where k = operation, i = job, t_{ik} = operation processing time

2.4 Problem statement

Determination of integrated schedule of tool and job which gives optimal or near optimal makespan, considering that n jobs are processed on m identical machines with T number of tool types are shared for many operations, in a flexible manufacturing system.

3 Proposed heuristics

MGA heuristic is used to find the optimal or near optimal solution for the proposed problem. The above algorithm uses a new algorithm ACT_SCH which is developed by modifying GT algorithm, for generating active feasible joint operation-tool schedule. The new algorithm ACT_SCH is described as below:

3.1 ACT_SCH algorithm for generating active feasible schedule

GT algorithm is basically used for generating active feasible schedule for a job shop scheduling. GT algorithm is a mathematical version of procedure for drawing Gantt

chart to arrive at active feasible schedule. The proposed new algorithm ACT_SCH is developed by using the similar procedure to GT algorithm. Whenever two or more jobs require the same tool at the same time (job conflict), the tool is assigned to a job that has minimum datum time (DT) for the waiting operation. Through this job conflict resolving policy, the waiting time of other jobs for the same tool reduces and hence an active schedule with lower makespan is obtained. The procedural step of ACT_SCH is explained in following Sect. 3.1.1; the flowchart is shown in Fig. 2 and illustrated in Sect. 3.1.2.

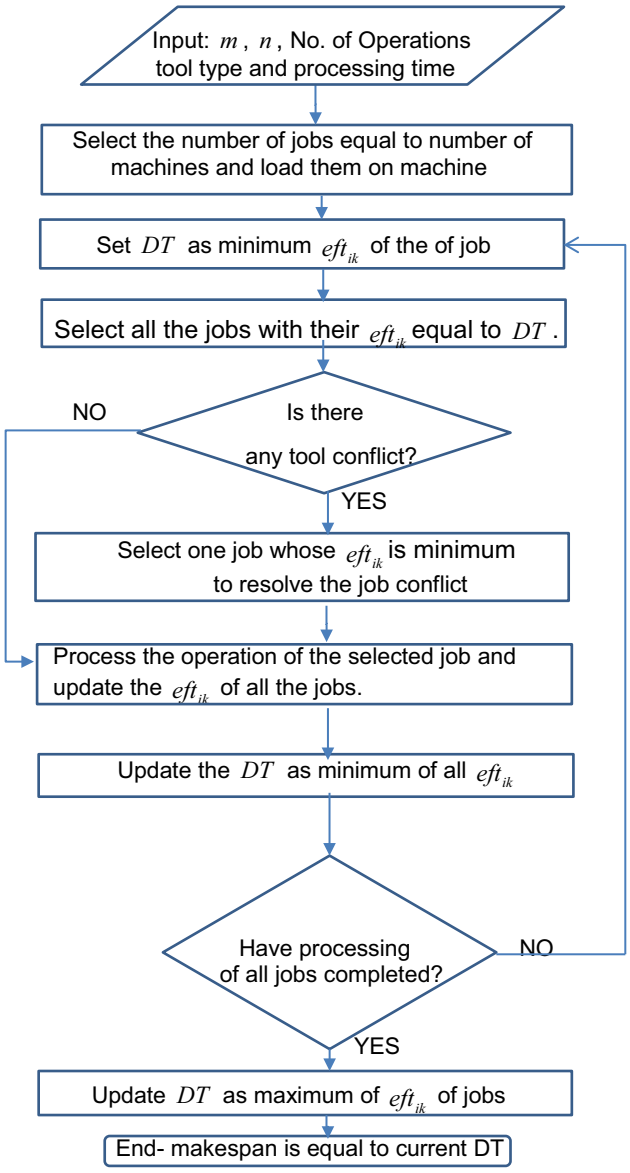


Fig. 2 Flow chart for ACT_SCH algorithm

3.1.1 *Procedural steps of ACT_SCH*

Step 1 Construct a table having columns one for each machine, job assigned (JA), datum time and processed operation of a job (POJ). The column for each machine is further subdivided for each tool type

Step 2 Select randomly the number of jobs equal to the number of machines and assign one job to one machine arbitrarily. Enter a value equal to sum of earliest start time (est_{ik}) and processing time (t_{ik}) in the first line of the table in appropriate block of machine under appropriate tool. The total of the two values gives earliest finish time (eft_{ik}) of the immediately waiting operation of the assigned job. For the first operation of first job assigned on a machine the value of est_{ik} is taken to be zero

Step 3 Enter the value of DT equal to the smallest of the eft_{ik} entries, in the appropriate column

Step 4 Select the jobs whose eft_{ik} matches with the current datum time. In case there are more than one jobs whose eft_{ik} matches with current datum time; select one job arbitrarily

Step 5 Check for the job conflict? If yes, go tostep 6 else go to step 7

Step 6 List the operations of all the jobs contending for the same tool. Select a job whose eft_{ik} is minimum and earmark its operation. Assign the tool to the earmarked operation of the job

Step7 Process the earmarked operation along with its machine and tool. Enter the job number and operation completed in the format of 'i_k' in the column "*P.O.J*"

Step 8 Update the value of eft_{ik} of next waiting operation as sum of earliest finishing time of immediate processed operation and processing time of the waiting operation. Also, update the eft_{ik} of the contending operation of other jobs (who are not assigned with the required tool during conflict resolution) as sum of earliest finish time of the operation of the job assigned with the tool plus its processing time

Step 9 Update the datum time as minimum of earliest finishing time. Check, if all the operations of all the jobs are scheduled. If yes, go to step 10, else go to step 4

Step 10 Update the DT as largest entry of earliest finishing time. The table gives the active feasible schedule and DT represents the makespan time

From the table the est_{ik} and eft_{ik} of each job can be read and subsequently the Gantt chart can be drawn.

3.1.2 Numerical example

A numerical example of scheduling problem of three jobs on two machines with three tool types is taken to explain the detailed procedure of the proposed algorithm (ACT_SCH). The sequence of jobs is taken as 2–1–3. The problem is given in the Table 1 and the procedure for generation of active feasible schedule is shown in the Table 2.

In row number 4 in Table 2, the operation number 3 of job number 2 is waiting on machine number 1 and operation number 2 of job number 1 is waiting on machine 2 to be processed and both require same tool type 1. This is a condition of conflict of job (COJ) and is resolved by selecting job 1 to be supplied with the tool as the earliest finish time of its operation is the least among all the contending jobs. Now, the operation 3 of job 2 has to wait till the tool type 1 gets free from the assigned job.

Table 1 Job-operation-tool matrix

Job no.	Operation no. (i)		
J	1	2	3
1	4 (2)[a]	10 (1)	8 (3)
2	6 (3)	5 (2)	6 (1)
3	7 (1)	6 (3)	8 (2)

[a] The value in parenthesis represents tool type for the processing of an operation

3.2 Modified genetic algorithm (MGA)

Genetic algorithms are stochastic search techniques that rely on the process of natural selection (Goldberg 1989). To start with simple GA, A set of random solutions called initial population, is generated. Each chromosome in this population contains information about a possible solution. The fitness of each chromosome is calculated based on the objective function. The chromosomes with higher fitness value are pooled in a group and are used for crossover. Mutation operator is applied to the offspring obtained through crossover, in order to maintain diversity of solution. Select the number of parents and offspring with higher fitness value, equal to the population size for next round of breeding and mutation. The process is repeated until optimal or near optimal solution is achieved.

It is important to control the distribution of offspring in comparison to their parents in order to improve the performance of GAs. The narrow distributions of offspring lead to premature converge of the algorithm, whereas wider distributions of offspring result in the algorithm taking a

Table 2 Active schedule generation procedure

Machine number						Job Processed		DT	COJ	P.O.J
1			2			Machine 1	Machine 2			
Tool type			Tool type							
1	2	3	1	2	3					
		0 + 6 6		0 + 4 4*		2	1	4	1	1_1
		0 + 6 6*	4 + 10 14			2	1	6	3	2_1
	6 + 5 11*		4 + 10 14			2	1	11	3	2_2
11 + 6 19			4 + 10 14*			2	1	14	2.1	1_2
14 + 6 20*					14 + 8 22	2	1	20	2	2_3
20 + 7 27					14 + 8 22*	3	1	22	1	1_3
20 + 7 27*						3	–	27	3	3_1
		27 + 6 33*				3	–	33	3	3_2
	33 + 8 41*					3	–	41	3	3_3

*Jobs selected for processing

long time to converge to the optimal solution. In this context, the research paper proposes a modified GA with three-parent crossover and a diversity operator. Three parents crossover produces three offspring by linear combination of three parents to improve the fitness and diversity of the offspring. The mutation operator further diversifies the generated offspring.

Initial population of chromosomes of size IP equal to $(n^2 - n)/2$ is generated. Each chromosome in this population contains numbers from 1 to j which are randomly generated by permutation. The structure of a chromosome containing five jobs is shown in the Fig. 3.

Each number in the chromosome signifies a job number and the placement of numbers in the chromosome from left to right represents their sequence of assignment on *m* machines. For instance, if there are 2 machines then first two jobs; 5 and 3 are assigned to machine 1 and 2 respectively. The next job in the sequence is assigned to a machine which becomes free after processing the previous job.

From the generated initial population *s* best solutions are selected and stored in an archive pool A. Now using tournament selection (with size 2), the chromosomes are selected and stored in the selection pool B with a pool size of 3**IP*. Three consecutive parents from the selection pool are selected and are arranged in increasing value of fitness i.e. $f(P_1) \leq f(P_2) \leq f(P_3)$. Any duplicity of the chromosomes is removed by replacing the unwanted chromosomes with a random chromosome from the selection pool. The crossover among these three chromosomes P_1, P_2 and P_3 is performed to generate three offspring. The procedure to generate offspring O_1 is explained in Fig. 4a, b. The whole process of crossover operation is divided into two steps. In the first step, as illustrated in Fig. 4a, parent chromosome P_2 crossover with P_3 using the randomly generated crossover vector x_1, to produce intermediate offspring IM_1. In this step, firstly the resultant parent P_2x_1 is obtained by replacing all non-zero elements of x_1 by the corresponding elements of parent P_2. Then all the zero elements in the resultant chromosome are replaced by the corresponding elements of second parent P_3 in order to generate intermediate offspring IM_1. In case, if any of the replacement of zero elements results in repeat of numbers in IM_1, then it has to be replaced by a non-repeating element of the first chromosome i.e. P_2. As shown in Fig. 4a the element 3 of chromosome P_3 cannot replace 0 element in P_2x_1, as number 3 already exist in the resultant chromosome. Hence, the first element of chromosome P_2 corresponding to zero element i.e. 4 is considered for the replacement. But this number would also cause duplicity, and therefore the next element in P_2 corresponding to 0 i.e. 5, is considered. Since it does not create duplicate entry in the resultant chromosome, therefore number 5 is used instead of number 3 to replace the 0 element in the resultant chromosome P_2x_1. The proposed method of crossover results into a feasible intermediate offspring IM_1. In the second step, third parent P_1 crossover with IM_1 as illustrated in Fig. 4b, to generate offspring O_1.

To obtain offspring O_2 and O_3 first of all, intermediate offspring IM_2 and IM_3 are generated by crossover of P_2

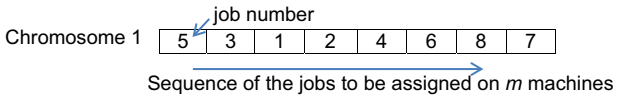


Fig. 3 Structure of a chromosome

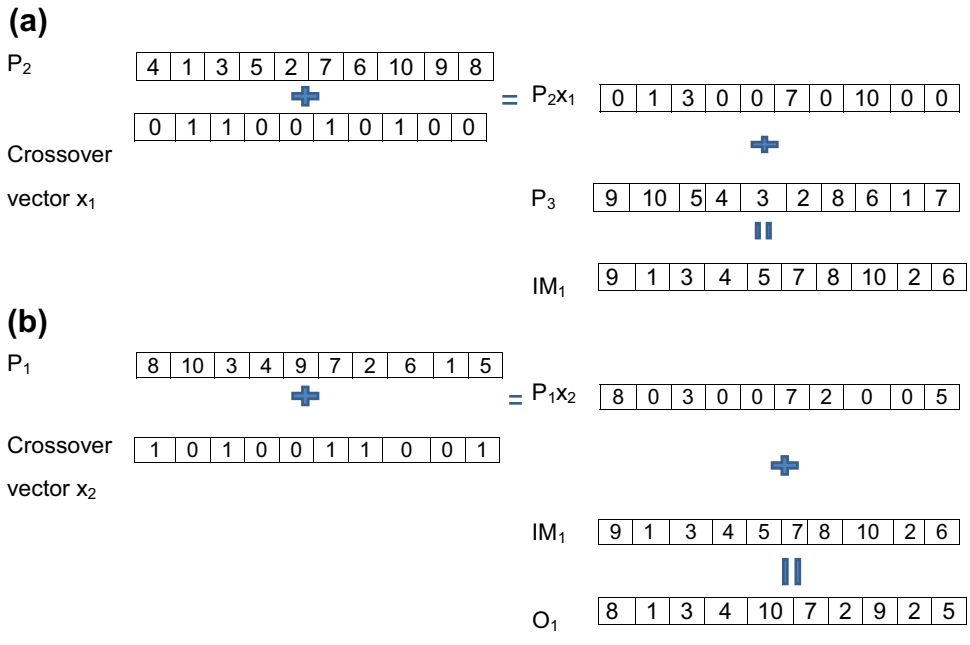


Fig. 4 **a** Step-I: parent chromosomes crossover using crossover vector x_1 to generate intermediate offspring IM_1. **b** Step-II: IM_1 and third parent chromosome P_1 crossover using crossover vector x_2 to generate final offspring O_1

with P_3 and P_3 with P_1 respectively, subsequently P_1 crossover with IM_2 and P_2 crossover with IM_3 to generate offspring O_2 and O_3 respectively.

In three parent crossover mechanism, it is observed that offspring O_1 and O_3 move towards better fitness while offspring O_3 diversify the population.

A mutation operator is used to diversify the population to help the algorithm to escape from local minima. In this, randomly selected two elements of a selected offspring are exchanged. The probability of mutation operation is taken as p. The offspring and archive pool individuals are then merged and the best s individuals are selected as new population for the next generation.

Various parameters for the MGA, like $s = 0.5 * IP$, $p = 0.05$, $MAXGEN = 100$ and crossover probability $= 1$ are decided based on the trial runs of modified Genetic Algorithm on five different scheduling problems from the literature (Prabaharan et al. 2006) for each case of $m = 2$, $m = 3$ and $m = 4$. The results are compared with the optimum value obtained by complete enumeration of the problems. It is observed that optimum results are obtained for 43.33% of the problems and 50% of the problem shows deviation <8% from optimum value.

The procedural step of Modified GA is explained in following Sect. 3.2.1; the flowchart is shown in Fig. 5.

3.2.1 Steps of modified GA

Step 1 Set $p = 0.05$, $IP = (n^2 - n)/2$, $MAXGEN = 100$ and $s = 0.5 * IP$

Step 2 An initial population of size IP is generated by random permutation of number of jobs j

Step 3 An archive A is filled with s best solutions

Step 4 Using the tournament selection process with size $TC = 2$, a selection pool B is filled with selected chromosomes with size $3IP$

Step 5 Select three consecutive chromosomes from the selection pool and arrange them so that $f(P_1) \leq f(P_2) \leq f(P_3)$. Replace any duplicate chromosome with randomly selected chromosome from the selection pool

Step 6 Crossover of three parents is performed to obtain three offspring

Step 7 Mutation operation of the selected offspring is performed

Step 8 Generate the active feasible schedule from the sequence of jobs in chromosomes by using ACT_SCH algorithm and calculate the makespan time

Step 9 Combine chromosomes from archive pool and offspring. Select the best individuals to make a new population of size IP

Step 10 Check, if termination criterion is met? If yes go to step 12 Otherwise go to step 2

Step 11 Stop and note down the optimal sequence and corresponding optimal makespan.

4 Results and discussions

To obtain the test results for makespan time for the proposed joint scheduling problem a set of 20 problems, each for $m = 2$, $m = 3$ and $m = 4$, are taken from the literature (Prabaharan et al. 2006). The results obtained for makespan

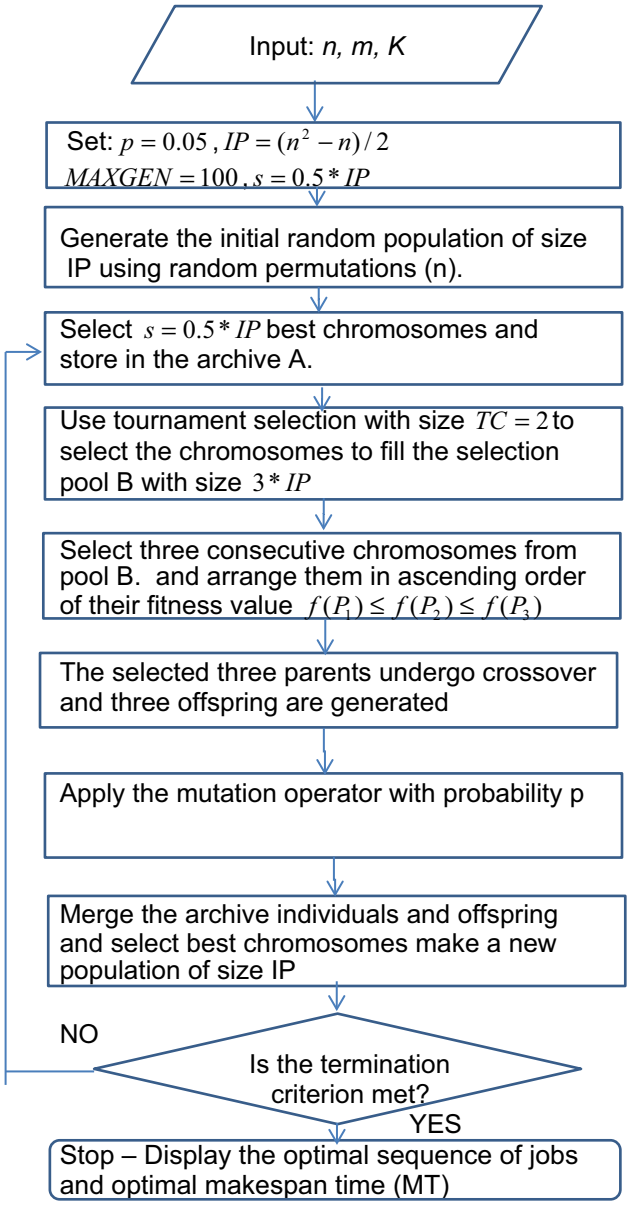


Fig. 5 Flowchart for MGA algorithm

time using proposed Modified GA are compared with the results reported in the literature in Tables 3, 4 and 5.

The results shown in fourth column of the tables represents the best value of makespan obtained by two techniques, namely PDRA and Simulated Annealing Algorithm (SAA), in the literature (Prabaharan et al. 2006), whereas the results reported by Udhayakumar and Kumanan (2010) are shown in the fifth column. The results obtained by the proposed MGA are shown in the last column. The best value of makespan time is marked with boldface. The computational time taken by the proposed MGA is shown in the Fig. 6.

The percentage difference from best (PDB) is calculated to measure the performance of the proposed algorithm in giving better results comparison to best results reported in the literature.

$$PDB = (MT_{best} - MT_{MGA}) * 100/MT_{best}$$

where, MT_{best} is the best makespan value by the other reported algorithms and MT_{MGA} is the value obtained by MGA.

The results show that the MGA algorithm gives better results for all the problems except for the problem number 3 and 14 for $m = 2$, and problem number 19 for $m = 4$. The average percentage difference calculated are 1.95% for $m = 2$, 5.25% for $m = 3$ and 7.29% for $m = 4$. The results also show that the average percentage difference increases with increase in number of machines which represents that the performance of proposed MGA increases with increase in problem size. The programming for ACT_SCH algorithm and MGA was developed using "MATLAB". The computational time taken by MGA is reasonable. The computational time increases exponentially with higher number of jobs. But the higher time does not delay the implementation of solution as the scheduling problems can be solved offline prior to the actual loading of the batch of jobs.

5 Conclusions

The application of MGA to the FMS loading and scheduling problem is explored in this paper. The FMS environment consists of several identical machines without tool magazines. A CTM stores all types of tools and shares the tools among the machines through the tool transport system. The Modified GA uses the principle of evolution to find the optimal solution of the scheduling problem. The three parent crossover generates better fitness and diversified offspring. Use of diversity operator helps the algorithm to avoid local minima and premature convergence.

In the proposed MGA, initial population is generated randomly and the best s individuals based on fitness values are stored in archive pool. Tournament selection with size 2 was used to select the individuals of size $3 * IP$ for the reproduction. Three consecutive individuals are selected

Table 3 Comparison of results obtained for m = 2

Problem no.	No. of jobs	No. of operations	Best solution PDRA and SAA	Solution by ACO	Proposed MGA algorithm
1	3	3	**38**	**38**	**38**
2	4	3	**99**	**99**	**99**
3	5	4	111	**103**	107
4	6	4	131	128	**126**
5	7	5	212	210	**208**
6	8	7	300	300	**293**
7	9	5	242	243	**208**
8	10	6	402	394	**385**
9	11	5	297	294	**282**
10	12	4	257	265	**248**
11	13	5	703	693	**669**
12	14	6	542	540	**525**
13	15	7	339	340	**334**
14	16	8	332	**328**	332
15	17	8	349	350	**344**
16	18	8	399	419	**397**
17	19	9	482	476	**468**
18	20	9	532	527	**524**
19	25	9	643	643	**635**
20	30	9	797	813	**795**

Table 4 Comparison of results obtained for m = 3

Problem no.	No. of operations	PDRA and SAA	Best solution	Solution by ACO	Proposed MGA algorithm
1	3	3	**25**	**25**	**25**
2	4	3	**90**	91	**90**
3	5	4	83	83	**78**
4	6	4	102	100	**95**
5	7	5	168	160	**153**
6	8	7	227	221	**216**
7	9	5	200	193	**168**
8	10	6	299	290	**280**
9	11	5	220	219	**202**
10	12	4	203	204	**182**
11	13	5	540	524	**475**
12	14	6	402	403	**367**
13	15	7	247	239	**233**
14	16	8	238	256	**232**
15	17	8	253	246	**242**
16	18	8	291	320	**272**
17	19	9	344	338	**324**
18	20	9	392	387	**364**
19	25	9	450	**441**	445
20	30	9	583	614	**556**

Table 5 Comparison of results obtained for m = 4

Problem no.	No of jobs	No. of operations	Best solution PDRA and SAA	Solution by ACO	Proposed MGA algorithm
1	3	3	–	–	–
2	4	3	87	87	**70**
3	5	4	78	79	**73**
4	6	4	90	87	**85**
5	7	5	142	136	**135**
6	8	7	192	181	**168**
7	9	5	158	149	**146**
8	10	6	270	258	**232**
9	11	5	196	182	**164**
10	12	4	187	187	**157**
11	13	5	457	448	**407**
12	14	6	328	318	**308**
13	15	7	210	198	**191**
14	16	8	200	189	**183**
15	17	8	217	203	**189**
16	18	8	247	238	**217**
17	19	9	271	259	**255**
18	20	9	322	316	**288**
19	25	9	366	**351**	355
20	30	9	464	452	**446**

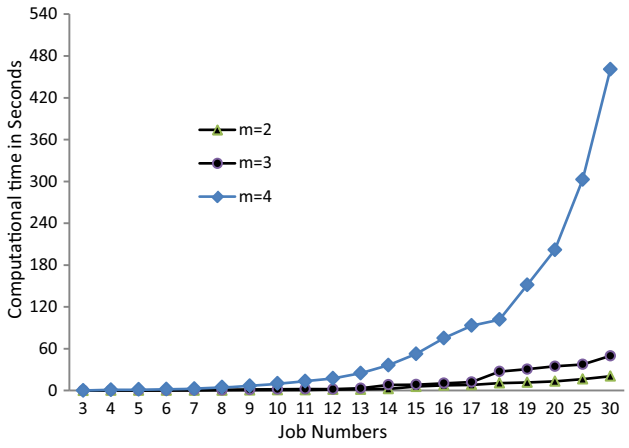


Fig. 6 Computational time taken by MGA

for reproduction by using the crossover operator. The mutation operator is then used to diversify the offspring. The offspring and archive pool individuals are then merged and the best s individuals are selected as new population for the next generation.

The ACT_SCH algorithm is used to generate the active feasible schedule. The results obtained for the makespan by the MGA are compared with the results reported in the literature and it was found that the proposed MGA algorithm gives optimal or near optimal results in almost all the test problems. The computation time taken by the MGA for $m = 2$ and $m = 3$ cases is comparable to that reported for ACO and SAA in the literature, but the same is higher for $m = 4$ case. Since the scheduling problems can be solved offline before the actual production starts, higher computational time does not delay the implementation of solution.

References

Agnetis A, Alfieri A, Brandimarte P, Prinsecchi P (1997) Joint job/tool scheduling in FMC with no on-board tool magazine. Comput Integr Manuf Syst 10(1):61–68

Balin S (2011) Non-identical parallel machine scheduling using genetic algorithm. Expert Syst Appl 38(6):6814–6821

Brucker P (1995) Scheduling algorithm, 1st edn. Springer, Berlin

Chan FTS, Chung SH, Chan LY (2008) An introduction of dominant genes in genetic algorithm for FMS. Int J Prod Res 46(16):4369–4389

Elsayed SM, Sarkar SA, Essam DL (2014) A new genetic algorithm for solving optimization problems. Eng Appl Artif Intell 27:57–69

Fathi Y, Barnette KW (2002) Heuristic procedure for parallel machine problem with tool switches. Int J Prod Res 40(1):151–164

French S (1982) Sequencing and scheduling, 5th edn. Hardwood, London

Gang X, Wu Z (2004) Deadlock-free scheduling strategy for automated production cell. IEEE Trans Syst Man Cybern Part A Syst Hum 34(1):113–122

Giffler B, Thompson GL (1960) Algorithms for solving production scheduling problems. Int J Oper Res 8:487–503

Gnanavel BA, Jerald J, NoorulHaq A, Muthu Luxmi M, Vigneswaralu TP (2010) Scheduling of machines and automated guided vehicles in FMS using differential evolution. Int J Prod Res 48(16):4683–4699

Godinho FM, Barco CF, Neto RFT (2014) Using genetic algorithms to solve scheduling problems on flexible manufacturing systems (FMS): a literature survey, classification and analysis. Flex Serv Manuf J 26(3):408–431

Goldberg DE (1989) Genetic algorithm in search, optimization, and machine learning. Addition-Wesley, Boston, MA

Gray AE, Seidmann A, Stecke KE (1993) A synthesis of decision models for tool management in automated manufacturing. Manag Sci 39:549–567

Holland HH (1975) Adaption in natural and artificial systems. University of Michigan Press, Detroit, MI

Hsu T, Korbas O, Dupas R, Gonclaves G (2008) Cyclic scheduling of FMS: modeling and evolutionary solving approach. Eur J Oper Res 191:464–484

Jun HB, Kim YD, Suh HW (1999) Heuristics for a tool provisioning problem in a flexible manufacturing system with an automatic tool transporter. IEEE Trans Robot Autom 15(3):488–496

Kaplanoğlu V (2016) An object-oriented approach for multi-objective flexible job-shop scheduling problem. Expert Syst Appl 45:71–84

Keung KW, Ip WH, Yuen D (2003) An intelligent hierarchical workstation control model for FMS. J Mater Process Technol 139:134–139

Kim YK, Kim JY, Shin KS (2007) An asymmetric multileveled symbiotic evolutionary algorithm for integrated FMS scheduling. J Intell Manuf 18(6):631–645

Mukhopadhyay SK, Nandi PK (1999) Solving tool allocation problem in flexible manufacturing system. J Ind Eng (India) 80:16–20

Nasciment MA (1993) Giffler and Thompson algorithm for job shop scheduling is still good for flexible manufacturing systems. J Oper Res Soc 44(5):521–524

Ponnambalam SG, Aravindan P, Srinivasa Rao P (2001) Comparative evaluation of genetic algorithms for job-shop scheduling. Prod Plan Control 12(6):560–574

Prabaharan T, Nakkeeran PR, Jawahar N (2006) Sequencing and scheduling of job and tool in a flexible manufacturing cell. Int J Adv Manuf Technol 29(7–8):729–745

Reddy BSP, Rao CSP (2006) A hybrid multi-objective GA for simultaneous scheduling of machines and AGVs in FMS. Int J Adv Manuf Technol 31:602–613

Roh HK, Kim YD (1997) Due date based loading and scheduling methods for a flexible manufacturing system with an automatic tool transporter. Int J Prod Res 35(11):2989–3003

Selim AM, Ozkan S (2001) Integrated scheduling and tool management in flexible manufacturing systems. Int J Prod Res 39(12):2697–2722

Selim AM, Siraceddin O (1999) Joint lot sizing and tool management in a CNC environment. Comput Ind 40:61–75

Shankar S, Ponnamabalam SG, Gurumarinuthu M (2005) Scheduling flexible manufacturing systems using parallelization of multi-objective evolutionary algorithms. Int J Adv Manuf Technol 30:279–285

Stecke KE (1983) Formulation and solution of non-linear integer production planning problem for flexible manufacturing system. Manag Sci 29(3):273–288

Turkcan A, SelimAkturk M, Storer RH (2007) Due date and cost-based FMS loading, scheduling and tool management. Int J Prod Res 45(5):1183–1213

Udhayakumar P, Kumanan S (2010) Sequencing and scheduling of job and tool in a flexible manufacturing system using ant colony optimization algorithm. Int J Adv Manuf Technol 50:1075–1084

Veeramani D, Upton DM, Barash MM (1992) Cutting tool management in computer integrated manufacturing. Int J Flex Manuf Syst 4(3–4):237–265

Wu MC, Lin CS, Lin CH, Chen CF (2017) Effects of different chromosome representations in developing genetic algorithms to solve DFJS scheduling problems. Comput Oper Res 80:101–112

Wireless Netw
DOI 10.1007/s11276-017-1533-y

Intelligent selection of threshold in covariance-based spectrum sensing for cognitive radio networks

Chhagan Charan[1] · **Rajoo Pandey**[1]

© Springer Science+Business Media New York 2017

Abstract The radio spectrum sensing has been an important issue of research in cognitive radio networks over the last decade and the appropriate selection of threshold plays a crucial role in the process of spectrum sensing. The conventional channel sensing methods generally employ a fixed threshold, which is either based on the principle of constant false alarm rate (CFAR) or the principle of constant detection rate (CDR). The sensing performance of these schemes degrades under low signal to noise ratio (SNR) and noise uncertainty. The problem of noise uncertainty occurring in energy detection (ED) based spectrum sensing method can be overcome by using covariance-based spectrum sensing scheme. However, the performance of covariance based spectrum sensing degrades at low SNR. This paper proposes a covariance-based channel sensing method, where the adaptive threshold is selected in an intelligent manner to minimize the probability of error with sufficient protection to primary user (PU). First, an adaptive threshold is derived by considering both probability of detection and probability of false alarm such that the total decision error probability is minimized. This adaptive threshold is then considered along with two other thresholds based on CFAR and CDR schemes, for the final selection of threshold such that the protection to PU is maximized. The proposed approach also provides the minimum number of samples required for reliable spectrum sensing. As shown by the simulation results, the proposed scheme exhibits better detection performance compared to ED based schemes as well as the existing covariance-based detection method in terms of probability of detection and probability of decision error.

Keywords Cognitive radio · Spectrum sensing · Covariance matrices · Intelligent selection of threshold · Adaptive threshold · Low SNR

1 Introduction

Over the last decade, the exponential growth of wireless communication technologies and surge of data services have given rise to an ever-increasing demand for spectrum resources. This has resulted in scarcity of frequency spectrum, which is a limited natural resource. Interestingly, the recent study of spectrum utilization carried out by Federal Communication Commission (FCC) has indicated that most of the spectrum is underutilized in vast temporal and geographical dimensions due to the fixed spectrum allocation methodology [1, 2]. Therefore, to resolve the conflict between spectrum underutilization and spectrum scarcity problem, the "Cognitive radio" (CR) technology is considered as a viable option [3, 4]. The cognitive radio has the ability to sense the huge swath of spectrum in external radio frequency environment and adapt itself by changing its communication parameters such as transmission power, carrier frequency and bandwidth accordingly [3, 5]. In CR, the unlicensed secondary user (SU) borrows the spectrum from licensed primary user (PU) or shares the spectrum with PU networks to increase the utilization of unused spectrum without interfering with the PU [3–5]. To ensure this, the SU has to dynamically detect the presence of licenced users [4, 5].

✉ Chhagan Charan
chhagan.charan@nitkkr.ac.in

Rajoo Pandey
rajoo_pandey@nitkkr.ac.in

[1] Department of Electronics and Communication Engineering, National Institute of Technology, Kurukshetra, India

Published online: 29 May 2017

The performance of a CR networks heavily depends on its ability to sense the frequency spectrum. There are various spectrum sensing techniques available in the literature such as energy detection (ED) [6–11], matched filter detection (MF) [12, 13], cyclostationary feature detection [14, 15], maximum–minimum eigenvalue detection [16, 17] and covariance-based detection [18–23] methods.

There are three different paradigms for spectrum sharing in the CR network, namely underlay, overlay and interweave [24]. In the underlay approach, the SU can always access the licensed spectrum by transmitting the signal at a very low power level so that the interference to the PU remains below the prescribed limit [25, 26]. In the overlay approach, however, SU also helps in PU's transmission while accessing the spectrum. In this mode, PU shares the information about signal codebook and message with the SU [27, 28]. In the interweave approach, the SU exploits the unused frequency spectrum band of the licensed PU in an opportunistic fashion [27]. The covariance based spectrum sensing scheme proposed in this paper can be used with overlay and interweave approaches of spectrum sharing for cognitive radio networks.

Many spectrum-sensing methods require the information about the spectrum utilization factor of licensed user [9, 23]. Therefore, spectrum occupancy measurements are very useful in CR system design and can be utilized to increase spectrum sensing accuracy. In the literature several spectrum occupancy models are studied to estimate the behavior of PU's spectrum utilization [29–34]. The covariance-based spectrum sensing approach proposed here also exploits the information about SNR and the spectrum occupancy by PU to obtain more reliable sensing performance.

The rest of the paper is organized as follows. Works related to spectrum sensing are reviewed in Sect. 2, and related issues are discussed. The spectrum sensing model and the conventional methods based on ED and covariance matrix are briefly described in Sect. 3. The proposed covariance-based adaptive threshold spectrum-sensing method and scheme for intelligent selection of threshold are presented in Sect. 4. The simulation results of the proposed scheme and a comparison with ED and existing covariance-based approaches are presented in Sect. 5. Finally, overall findings of this study are concluded in Sect. 6.

2 Related work

The spectrum sensing is a vital component in CR networks. Recently, various spectrum-sensing methods have been proposed in the literature [7–23]. Among these methods, the energy detection [7–11] is widely used as it does not require prior knowledge of PU's signal characteristics and is also easy to implement. However, as shown by several studies [35, 36], its performance severely degrades at low SNR and is not robust to noise uncertainty. In order to improve the performance at low SNR, authors discussed a spectrum-sensing algorithm in [10], which is based on three consecutive events. This algorithm, named as three-event ED (3EED), takes the decision in one sensing event, considering also the event immediately before and the one immediately after it. It exploits the knowledge of duty cycle of the PU's activity for tracking the changes in the PU's state. However, this scheme is not able to address the problem of noise uncertainty. Also, these methods require an estimate of noise power [7, 9, 10], which in practice, is very challenging to obtain [35]. The matched filtering (MF)-based scheme [12, 13] can perform well at low SNR and with noise uncertainty but requires prior information of licensed user. Similarly, the cyclostationary detection method [14, 15] needs the knowledge of the PU's cyclic frequency. Therefore, both MF and cyclostationary-based detection methods are not suitable for blind spectrum sensing. To overcome the problem of noise uncertainty, the maximum eigenvalue detection method [16, 17] and covariance-based method [18–23] have been proposed. The covariance-based spectrum sensing methods use the statistical covariance of signal and noise, which are usually different. This property is exploited to decide whether PU is present or not [18–22]. Among other schemes based on this principle, the weighted-covariance-based spectrum sensing [22] exploits the spatial/temporal correlation of PU signal. The graph-theory based detection scheme is discussed in [23], which maps received data to the graph-related data sheets and examines the corresponding adjacency matrix and associated matrix converting the problem of spectrum sensing into classical problem of graph theory. The performance of all these existing covariance-based spectrum sensing methods degrade at low SNR, as they do not consider other factors such as signal to noise ratio and spectrum utilization ratio of PU.

In every spectrum sensing scheme, the decision about the presence or absence of a licensed user is based on a threshold, which is computed by using some test statistic. Most of the spectrum sensing techniques consider a constant false alarm rate to obtain a suitable value of threshold for decision making. The performance of spectrum sensing greatly depends upon the value of threshold used in the sensing scheme. The conventional spectrum sensing methods use a fixed threshold [7, 10, 18, 21], which does not ensure sufficient protection to PU. Also, the performance of these methods severely deteriorates at low SNR. It is very difficult to achieve pre specified values of both probability of detection and probability of false alarm with a single

value of threshold [37]. To address this issue a number of adaptive threshold-based techniques have been introduced [8, 9]. In [25] an optimal threshold is obtained through tradeoff between probability of detection and probability of false alarm by employing a weighted function. Sobron et al. [8] have discussed an adaptive spectrum sensing scheme using a cost function that consists of the aggregate information about absence or presence of PUs.

These schemes exhibit an improved performance at low SNR due to adaptive nature of the threshold employed. However, in case of noise uncertainty, the performance severely deteriorates [36]. Although the conventional covariance-based method can effectively address the problem of noise uncertainty, its performance at low SNR is not good [20, 21]. Therefore, the work presented in this paper attempts to develop a spectrum sensing algorithm, which can effectively overcome the challenges of noise uncertainty as well as that of low SNR. To achieve this goal, the proposed method named as covariance-based adaptive threshold (CAT) utilizes the information about PU's activity and SNR. The spectrum utilization factor of licensed user is obtained using the statistical spectrum occupancy model presented in [34] that employs first and second order parameters derived from statistical characteristic of frequency spectrum, and predicts spectrum occupancy with high accuracy. Further, the overall probability of decision error is minimized to obtain the adaptive threshold, which ensures better performance at low SNR. The proposed scheme is also robust against noise uncertainty as the test statistics used in the scheme are based on signal covariance like other conventional covariance-based spectrum sensing scheme.

The proposed work of CAT scheme is further extended to provide sufficient protection to PU along with minimization of decision error. The new scheme, named as covariance-based adaptive threshold with intelligent selection (CAT-IS), intelligently selects the decision threshold from a set of three values of threshold obtained using three different criteria, namely constant false alarm rate (CFAR), constant detection rate (CDR) and the proposed CAT. The threshold obtained using CDR criterion is best suited to provide sufficient protection to licensed user. The threshold obtained under CFAR criterion ensures the maximum utilization of spectrum by SU for given value of CFAR. The threshold obtained by CAT scheme achieves minimum probability of decision error. The proposed CAT-IS scheme varies the number of samples used in sensing to achieve the goal of minimization of decision error with sufficient protection to PU.

3 Spectrum sensing preliminaries

In this section, first of all, the commonly used spectrum sensing model is described. Then, a brief review of commonly used ED and covariance-based spectrum sensing technique is presented. Also, the notation adopted in this paper is described in Table 1.

3.1 Spectrum sensing model

Let the frequency band to be sensed by the SU be centred at frequency f_c having a bandwidth W. The received signal can be expressed as the following hypotheses [8, 13, 18]:

$$H_0 : x(n) = \eta(n) \tag{1}$$

$$H_1 : x(n) = s(n) + \eta(n) \tag{2}$$

where $x(n)$ is the sample of signal received by the SU, $s(n)$ is the sample of PU's transmitted signal with zero mean and variance σ_s^2, and $\eta(n)$ is additive white Gaussian noise (AWGN). The noise samples are assumed to be independent and identically distributed (IID) with zero mean and variance σ_η^2. Thus, the signal to noise ratio (SNR) is $\xi = \sigma_s^2 \big/ \sigma_\eta^2$.

The hypothesis H_0 indicates that the PU is absent whereas the hypothesis H_1 represents that PU is present. The performance of spectrum sensing can be measured in terms of probability of detection (P_d) i.e. a busy frequency band is detected to be busy, and probability of false alarm (P_f) that an idle frequency band is detected to be busy. To achieve the maximum throughput of SU, P_f should be as small as possible, whereas to provide protection to PU, P_d should be high [37].

3.2 Energy based spectrum sensing

In ED-based spectrum sensing, following test statistic is used [6–11]:

$$T(x) = \frac{1}{N}\sum_{n=1}^{N} |x(n)|^2 \tag{3}$$

where N is the number of samples used in test statistic.

The test statistic $T(x)$ of (3) has the Chi square distribution. However, when N is large, it can be approximated by Gaussian distribution using the central limit theorem [6, 7]. In conventional energy based spectrum sensing for a given value of probability of false alarm, the threshold λ can be expressed as described in [7, 24–28] by:

$$\lambda = \sigma_\eta^2 \left(\sqrt{\frac{2}{N}} Q^{-1}(P_f) + 1 \right) \tag{4}$$

Table 1 Notation adopted in this paper

Symbol	Definition
H_0	Null hypothesis indicating absence of PU
H_1	Alternate hypothesis indicating presence of PU
$x(n)$	Sample of signal received by the SU
$s(n)$	Sample of PU's transmitted signal
$\eta(n)$	Sample of additive white Gaussian noise
σ_η^2	Variance of noise
σ_s^2	Variance of PU's transmitted signal
ξ	Signal to noise ratio (SNR)
P_d	Probability of detection
P_f	Probability of false alarm
$T(x)$	Test statistic used for spectrum sensing
λ	Threshold in ED-based detection
N, N_s	Number of used sensing samples
$R_x(N_s)$	Autocorrelation matrix of received signal at SU
$R_s(N_s)$	Autocorrelation matrix signal transmitted by PU
$R_\eta(N_s)$	Autocorrelation matrix of noise
L	Smoothing factor
$r_{ij}(N_s)$	Element of matrix $R_x(N_s)$ at the *ith* row and *jth* column
$T_1(N_s)$, $T_2(N_s)$	Test statistics in covariance-based schemes
γ	Threshold for covariance-based detection
γ_{P_f}	Threshold for CFAD-based scheme
γ_{P_d}	Threshold for CDR-based scheme
Υ_L	Overall correlation strength among the consecutive L samples
P_e	Probability of decision error
α	Spectrum utilization factor of PU
γ_e	Threshold in CAT scheme
P_f^{th}	Maximum allowable value of P_f
P_d^{th}	Minimum allowable value of P_d
γ_f^*	Value of threshold used in CFAR-based scheme
γ_d^*	Value of threshold used in CDR-based scheme
γ^*	Value of threshold used in CAT-IS-based scheme
B	Noise uncertainty factor in decibel (dB)

where P_f is probability of false alarm and the Q-function is described as: $Q(z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$.

3.3 Covariance-based spectrum sensing method

The signal received at SU has the following covariance matrix [18–21]:

$$R_x(N_s) = \begin{bmatrix} g(0) & g(1) & \dots & g(L-1) \\ g^*(1) & g(0) & \dots & g(L-2) \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ g^*(L-1) & g^*(L-2) & \dots & g(0) \end{bmatrix} \quad (5)$$

where N_s denotes the number of samples used for computation of sample covariance matrix. The element $g(l)$ of the covariance matrix $R_x(N_s)$ is computed as:

$$g(l) = \frac{1}{N_s} \sum_{m=0}^{N_s-1} x(m) x^*(m-l) \quad (6)$$

with $l = 0, 1, \dots, L-1$; where L represents the smoothing factor [18, 20, 21].

The threshold used in the covariance absolute value (CAV) spectrum sensing scheme [18], is defined as:

$$\gamma = \frac{1 + (L-1)\sqrt{\frac{2}{N_s \pi}}}{1 - Q^{-1}(P_f)\sqrt{\frac{2}{N_s}}} \quad (7)$$

4 Proposed method for intelligent selection of threshold for covariance-based spectrum sensing

The performance of spectrum sensing in a cognitive radio system is characterized by P_d and P_f. The conventional spectrum sensing methods use a fixed threshold which can be based on constant false alarm rate (CFAR) or on constant detection rate (CDR). If the cognitive radio network is to be designed to ensure PU's safety, the CDR method is best suited and a high value of the target detection probability is chosen for threshold calculation. The higher the target P_d, higher the protection to PU from SU transmission. Whereas, if design priority of CR network is to maximize the spectrum efficiency of the secondary user, the CFAR method is used and the target P_f is kept very small. The lesser the value of target P_f, more the chances of spectrum band being utilized by the SU. In addition to above two techniques, threshold can also be obtained by minimizing the overall probability of error. In this paper, an adaptive threshold is derived and an algorithm of intelligent selection of threshold is presented. The adaptive threshold based on minimization of overall probability of error can take into account both P_d and P_f to ensure that the interest of both PU and SU are taken care of. The derivation of proposed adaptive threshold based on this criteria is presented in part *B* of this section. It has been observed that the value of P_d in this method degrades at low SNR.

In order to overcome this problem, an intelligent selection of threshold is proposed, which selects the optimum value of threshold from the three possible values of threshold based on above criteria. This method effectively maintain the desirable value of P_d with minimum value of overall probability of error.

4.1 Computation of thresholds for covariance-based CFAR and CDR schemes

Let the auto-correlation matrix of the signal received at SU be expressed as:

$$R_x(N_s) = R_s(N_s) + R_\eta(N_s) \tag{8}$$

where $R_x(N_s)$, $R_s(N_s)$ and $R_\eta(N_s)$ represent the autocorrelation matrices of the received signal at SU, signal transmitted by the licensed user, and noise, respectively.

It is assumed that the noise is additive white Gaussian with zero mean and σ_η^2 variance, hence Eq. (8) can be expressed as:

$$R_x(N_s) = R_s(N_s) + \sigma_\eta^2 I_L \tag{9}$$

where I_L is an identity matrix of size $L \times L$.

Then, if PU's transmitted signal $s(n)$ is absent i.e. $R_s(N_s) = 0$, the off-diagonal elements of $R_x(N_s)$ are all nulls. Whereas, if signal $s(n)$ is present and signal samples are correlated, $R_x(N_s)$ will have significant values of off-diagonal elements.

In this method following test statistics are computed:

$$T_1(N_s) = \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{L}\left|r_{ij}(N_s)\right| \tag{10}$$

and

$$T_2(N_s) = \frac{1}{L}\sum_{i=1}^{L}\left|r_{ij}(N_s)\right| \tag{11}$$

Here, $r_{ij}(N_s)$ represents the element of matrix $R_x(N_s)$ at the *ith* row and *jth* column. Therefore, if there is no signal $T_1(N_s)/T_2(N_s) = 1$; and in the presence of signal $T_1(N_s)/T_2(N_s) > 1$. Thus, the ratio $T_1(N_s)/T_2(N_s)$ forms the test statistic $T(N_s)$, to detect the presence of signal in CFAR and CDR based methods as explained below.

1. In CFAR based scheme, the threshold is computed for the specified value of P_f. The P_f under the hypothesis H_0, it can be expressed as:

$$P_f = P(T_1(N_s) \geq \gamma_{P_f} T_2(N_s)) \tag{12}$$

Now the associated threshold γ_{P_f} is obtained in [14] as:

$$\gamma_{P_f} = \frac{1 + (L-1)\sqrt{\frac{2}{N_s\pi}}}{1 - Q^{-1}(P_f)\sqrt{\frac{2}{N_s}}} \tag{13}$$

2. If the CR network is designed to guarantee the PU's safety against the interference of SU by using the CDR method with certain target detection probability, then under hypothesis H_1, it can be expressed as:

$$P_d = P(T_1(N_s) \geq \gamma_{P_d} T_2(N_s)) \tag{14}$$

For a very large number of samples N_s and a low SNR, probability of detection (P_d) can be obtained as shown in [18] by the following expression:

$$P_d = 1 - Q\left(\frac{\frac{1}{\gamma_{P_d}} + \frac{\Upsilon_L \xi}{\gamma_{P_d}(\xi+1)} - 1}{\sqrt{2/N_s}}\right) \tag{15}$$

where $\Upsilon_L \triangleq \frac{2}{L}\sum_{l=1}^{L-1}(L-1)|\alpha_l|$, is the overall correlation strength among the consecutive L samples and α_l is the normalized correlation among the signal samples, expressed as:

$$\alpha_l = E[s(n)s(n-l)]/\sigma_s^2.$$

For the present case, the expression (15) can be utilized to compute the threshold (γ_{P_d}) which can ensure a given probability of detection as given below:

$$\gamma_{P_d} = \frac{1+\frac{\Upsilon_L \xi}{(\xi+1)}}{\sqrt{\frac{N_s}{2}}Q^{-1}(1-P_d)+1} \tag{16}$$

4.2 Proposed adaptive threshold based on minimization of overall probability of error

The tradeoff between P_d and P_f can be exploited to derive an adaptive threshold for spectrum sensing which minimizes the overall probability of decision error (P_e). For this purpose, the following expression is used:

$$\min(P_e(\gamma)) = \min\{P(H_0)P_f + P(H_1)(1-P_d)\} \tag{17}$$

where the probabilities $P(H_0)$ and $P(H_1)$, respectively, indicate the absence and presence of PU. The argument in (17) includes two types of errors: (1) deciding H_1 in place of H_0 (2) deciding H_0 instead of H_1. The adaptive threshold can be obtained by minimizing the following objective function.

$$P_e(\gamma) = (1-\alpha)P_f + \alpha(1-P_d) \tag{18}$$

Substitution of P_f and P_d from (13) and (15) into (18) yields:

$$P_e(\gamma) = (1-\alpha)\left(1 - Q\left(\frac{\frac{1}{\gamma}\left(1+(L-1)\sqrt{\frac{2}{N_s\pi}}\right)-1}{\sqrt{\frac{2}{N_s}}}\right)\right) + \alpha Q\left(\frac{\frac{1}{\gamma}+\frac{\Upsilon_L\xi}{\gamma(\xi+1)}-1}{\sqrt{2/N_s}}\right) \tag{19}$$

Let $U = \frac{\frac{1}{\gamma}\left(1+(L-1)\sqrt{\frac{2}{N_s\pi}}\right)-1}{\sqrt{\frac{2}{N_s}}}$ and $V = \frac{\frac{1}{\gamma}+\frac{\Upsilon_L\xi}{\gamma(\xi+1)}-1}{\sqrt{2/N_s}}$ Then

$$P_e(\gamma) = \frac{(\alpha-1)}{\sqrt{\pi}}\int_{\frac{U}{\sqrt{2}}}^{\infty} e^{-z^2}dz + \frac{\alpha}{\sqrt{\pi}}\int_{\frac{V}{\sqrt{2}}}^{\infty} e^{-z^2}dz + 1 - \alpha \tag{20}$$

For a known value of α, the total decision error probability $P_e(\gamma)$ is a convex function of threshold γ.

The probability of error in (19) can be minimized by setting the derivative equal to zero, which leads to the following expression:

$$\begin{aligned}&\gamma^2 \ln\left[\frac{(1-\alpha)}{\alpha}\frac{1+(L-1)\sqrt{\frac{2}{N_s\pi}}}{1+\frac{\Upsilon_L\xi}{(\xi+1)}}\right]\frac{4}{N_s}\\ &+\gamma\left[-2(L-1)\sqrt{\frac{2}{N_s\pi}}+\frac{2\Upsilon_L\xi}{(\xi+1)}\right]\\ &-\left[(L-1)^2\frac{2}{N_s\pi}+2(L-1)\sqrt{\frac{2}{N_s\pi}}\right.\\ &\left.-\frac{\Upsilon_L^2\xi^2}{(\xi+1)^2}-\frac{2\Upsilon_L\xi}{(\xi+1)}\right]=0\end{aligned} \tag{21}$$

Equation (21) can be written as:

$$\gamma^2 A + \gamma B + C = 0 \tag{22}$$

where

$$\begin{aligned}A &= \ln\left[\frac{(1-\alpha)}{\alpha}\frac{1+(L-1)\sqrt{\frac{2}{N_s\pi}}}{1+\frac{\Upsilon_L\xi}{(\xi+1)}}\right]\frac{4}{N_s},\\ B &= \left[-2(L-1)\sqrt{\frac{2}{N_s\pi}}+\frac{2\gamma_L\xi}{(\xi+1)}\right] \quad \text{and}\\ C &= -\left[(L-1)^2\frac{2}{N_s\pi}+2(L-1)\sqrt{\frac{2}{N_s\pi}}\right.\\ &\left.-\frac{\gamma_L^2\xi^2}{(\xi+1)^2}-\frac{2\gamma_L\xi}{(\xi+1)}\right]\end{aligned} \tag{23}$$

Let

$$\begin{aligned}a &= \left[(L-1)\sqrt{\frac{2}{N_s\pi}}-\frac{\gamma_L\xi}{(\xi+1)}\right] \quad \text{and}\\ b &= \left[(L-1)\sqrt{\frac{2}{N_s\pi}}+\frac{\gamma_L\xi}{(\xi+1)}+2\right]\end{aligned} \tag{24}$$

Then, it is easily observed that: $B = -2a$ and $C = -ab$.

Now, Eq. (22) can be expressed as:

$$\gamma^2 A - 2a\gamma - ab = 0 \tag{25}$$

The solution of Eq. (25) can be given as:

$$\begin{aligned}\gamma_1 &= \frac{a}{A}\left[-1+\sqrt{\left(1+\frac{Ab}{a}\right)}\right] \quad \text{and}\\ \gamma_2 &= \frac{a}{A}\left[-1-\sqrt{\left(1+\frac{Ab}{a}\right)}\right]\end{aligned}$$

The value of γ, selected to represent the threshold is positive i.e. $\gamma = \gamma_1$ and is described by:

$$\gamma_e = \left[\frac{\left((L-1)\sqrt{\frac{2}{N_s\pi}} - \frac{\Upsilon_L\xi}{(\xi+1)}\right)}{\ln\left[\frac{(1-\alpha)}{\alpha}\frac{1+(L-1)\sqrt{\frac{2}{N_s\pi}}}{1+\frac{\Upsilon_L\xi}{(\xi+1)}}\right]\frac{4}{N_s}}\right] \left[-1+\left(1+\frac{\left((L-1)\sqrt{\frac{2}{N_s\pi}}+\frac{\Upsilon_L\xi}{(\xi+1)}+2\right)}{\left((L-1)\sqrt{\frac{2}{N_s\pi}}-\frac{\Upsilon_L\xi}{(\xi+1)}\right)}\right.\right. \left.\left.\ln\left[\frac{(1-\alpha)}{\alpha}\frac{1+(L-1)\sqrt{\frac{2}{N_s\pi}}}{1+\frac{\Upsilon_L\xi}{(\xi+1)}}\right]\frac{4}{N_s}\right)^{0.5}\right] \quad (26)$$

where Υ_L represents the overall correlation strength among the consecutive L samples.

The parameter Υ_L can be obtained by using the test statistics $T_1(N_s)$ and $T_2(N_s)$. Under hypothesis H_1, Υ_L can be obtained from off-diagonal elements of the $R_x(N_s)$. It can be observed that

$$\lim_{N_s\to\infty} E(T_1(N_s) - T_2(N_s)) = \frac{2\sigma_s^2}{L}\sum_{l=1}^{L-1}(L-1)|\alpha_l| \quad (27)$$

where α_l denotes the correlation among signal samples.

Therefore, the overall correlation strength among the consecutive L samples can be estimated at SU as:

$$\Upsilon_L = \frac{T_1(N) - T_2(N)}{SNR\sigma_\eta^2} \quad (28)$$

where Υ_L is computed by using N samples, which also includes the samples of the previous sampling intervals. Thus, $N > > N_s$.

4.3 Proposed scheme of intelligent selection of threshold

The performance of the scheme based on the adaptive threshold introduced in part *B* can be further improved by adopting an intelligent strategy for the selection of a threshold, which takes into consideration the strengths of methods based on CFAR and CDR principles. The proposed scheme ensures sufficient protection to PU as well as minimizes the probability of overall decision error. For reliable spectrum sensing, the threshold (γ) should be selected in such a way that probability of false alarm $P_f(\gamma)$ is below a given specified value P_f^{th} and the probability of detection $P_d(\gamma)$ is above a given specified value P_d^{th}. Since $P_d(\gamma)$ and $P_f(\gamma)$ both decrease with increase in threshold γ, the above requirements translate into the following conditions:

$$\gamma \le \gamma_d^* \quad \text{and} \quad \gamma \ge \gamma_f^* \quad (29)$$

where γ_d^* and γ_f^* denote values of threshold required to satisfy the given values of probability of detection (P_d^{th}) and probability of false alarm (P_f^{th}), respectively. That is,

$$P_d(\gamma_d^*) = P_d^{th} \quad \text{and} \quad P_f\left(\gamma_f^*\right) = P_f^{th} \quad (30)$$

If $\gamma_d^* \ge \gamma_f^*$, then any value of γ which lies in the interval $\left[\gamma_f^*, \gamma_d^*\right]$ can satisfy the requirements of probability of detection and probability of false alarm. If $\gamma_d^* \le \gamma_f^*$, then there is no feasible value of γ, which can satisfy both these requirements. However, in this case a suitable threshold can still be selected if the condition $\gamma_d^* \ge \gamma_e$ is satisfied. Under this condition the choice of threshold as γ_e, can ensure sufficient protection to PU and minimizes the probability of overall error, which in turn implies reduction of P_f also.

The process of selection of optimum threshold is shown in the in the following flow chart:

Thus, in the proposed scheme, first of all, the threshold γ_f^*, γ_d^* and γ_e are computed using (13), (16) and (26), respectively for a given sample size N_s. Then $\gamma_d^* \ge \gamma_f^*$ is true and γ_e lies in the interval $\left[\gamma_f^*, \gamma_d^*\right]$, the choice of threshold as γ_e will satisfy the given requirement of P_d^{th} and P_f^{th}. If the condition $\gamma_d^* \ge \gamma_f^*$ is not satisfied or if γ_e does not lie in the interval $\left[\gamma_f^*, \gamma_d^*\right]$ than the condition $\gamma_d^* \ge \gamma_e$ is evaluated and if found true, the threshold can still be selected as γ_e. It can be noted that in this case the given requirement of P_f^{th} is not satisfied. However, the protection of PU is ensured with minimum value of P_f through minimization of overall probability of decision error. When the condition $\gamma_d^* \ge \gamma_e$ is also not found true, there is a need to increase the sample size by including some new samples through re-sensing. For example, the number of sample may be increased to $2N_s$ by combining the samples of previous intervals. By increasing the sample size the conditions $\gamma_d^* \ge \gamma_f^*$; $\gamma_d^* \ge \gamma_e$ or both may be satisfied so that the selection of an appropriate threshold becomes feasible.

Therefore, for reliable spectrum sensing the condition $\gamma_d^* \ge \gamma_e$ should be satisfied, where both γ_d^* and γ_e are function of number of samples (N_s). The minimum number of samples required for reliable spectrum sensing is the value of N_s that satisfies $\gamma_d^* = \gamma_e$. That can be easily obtained using (15) and (26) by employing some numerical technique.

4.4 Proposed covariance-based spectrum sensing using Intelligent Selection of Threshold

The algorithm of the proposed scheme for spectrum sensing named as covariance based adaptive threshold with intelligent selection (CAT-IS) is summarized below:

Step 1 Choose the smoothing factor L and the number of samples (N_s)

Step 2 Compute the autocorrelation of received signal using Eq. (6) and form the sample covariance matrix using (5)

Step 3 Compute the test statistics

$$T_1(N_s) = \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{L}\left|r_{ij}(N_s)\right| \quad \text{and}$$

$$T_2(N_s) = \frac{1}{L}\sum_{i=1}^{L}\left|r_{ij}(N_s)\right|$$

where $r_{ij}(N_s)$ represent the elements of the sample covariance matrix $R_x(N_s)$

Step 4 Obtain the overall correlation strength (Υ_L) among the consecutive L samples by using (28) and, using a period of 30 sensing intervals

Step 5 Calculate the optimal decision threshold γ^* for specified spectrum utilization ratio α using flowchart given in Fig. 1

Step 6 Obtain the test statistic

$$T(N_s) = \frac{T_1(N_s)}{T_2(N_s)} \tag{31}$$

Step 7 Take the sensing decision as:

$$D = \begin{cases} 1 & T(N_s) \geq \gamma^* \\ 0 & T(N_s) < \gamma^* \end{cases} \tag{32}$$

Here, the sensing decision $D = 0$ implies that the PU is not present, whereas sensing decision $D = 1$ indicates that PU is present.

In general, the computational complexity of any covariance-based spectrum-sensing method is higher than that of ED-based method. It can be easily observed that in the covariance-based spectrum sensing scheme computation of autocorrelation matrix involves $LN_s + L^2$ multiplications and additions. However, in the proposed spectrum sensing approach when the condition $\gamma_d^* \geq \gamma_e$; is not satisfied then the number of samples is increased as:

$$N_s^{new} = N_s + \Delta N_s \tag{33}$$

where N_s^{new} is the updated number of samples after re-sensing and ΔN_s represents the increase in the number of samples. The covariance matrix of received signal is updated as:

$$R_x\left(N_s^{new}\right) = \theta R_x(N_s) + (1-\theta)R_x(\Delta N_s) \tag{34}$$

where $\theta = N_s/(N_s + \Delta N_s)$

In addition to the covariance matrix, the proposed scheme also compares the thresholds for optimal selection. However, the computational efforts required for this are not significant in comparison with efforts required for computation of covariance matrix. The computational complexity increases as $O(N_s)$ in both the existing covariance-based technique and the present approach.

It can also be noted that the number of samples N_s is kept fixed in [18], whereas in the present approach it varies with SNR. At low SNR, the present scheme requires larger number of samples than in the existing scheme. Whereas, at high SNR it is less than that used by the existing covariance-based method. Thus, the overall computational complexity of the proposed method is same as that of [18].

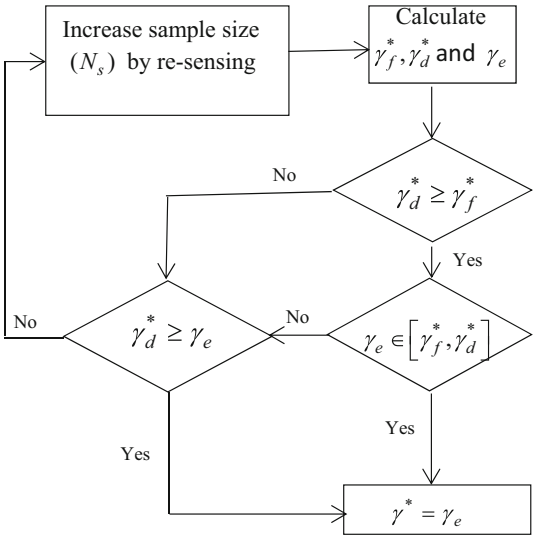


Fig. 1 Flow chart for the proposed scheme of intelligent selection of threshold

5 Simulation results and discussions

In this section, simulation results are presented to demonstrate the performance of the proposed covariance based adaptive threshold (CAT) spectrum sensing method and the covariance based adaptive threshold with intelligent selection (CAT-IS) spectrum sensing scheme. For comparison, existing schemes viz conventional ED-based spectrum sensing method, 3EED-based scheme and covariance absolute value (CAV)-based [18] spectrum sensing methods are also simulated.

The CR system model is simulated using MATLAB 8.1 and for performance analysis results are obtained by averaging 1000 Monte-Carlo runs. The statistical spectrum occupancy model presented in [34] is employed to estimate the PU's activity, which is varied in the range of 10–80%. The technique of [34], which estimates the utilization factor of PU, provides good accuracy ($\sim$91%) of spectrum prediction. The number of samples (N_s) in a sensing interval, used in the simulation is 15,000 and smoothing factor is chosen is $L = 5$. The maximum allowable

probability of false alarm and minimum allowable probability of detection is fixed at 0.1 and 0.9, respectively.

The variation of threshold value with SNR for proposed covariance-based adaptive threshold (CAT) scheme under different values of spectrum utilization factor (α) is shown in Fig. 2 for $N_s = 10,000$. From this figure, it is observed that the threshold varies with SNR to minimize the decision error probability whereas the threshold of existing CAV method is constant. Figure 2 is magnified into Fig. 3 to capture the variation of threshold at low SNR. It is noted that as the utilization ratio of PU increases, it requires smaller value of adaptive threshold to achieve the reliable spectrum detection, and vice versa. This is not possible in a fixed threshold-based scheme such as CAV method [18].

The variation of the thresholds γ_f^*, γ_d^* and γ_e with SNR is shown in Fig. 4 for $N_s = 25,000$. From this figure, it is observed that the value of threshold γ_e, in covariance-based adaptive threshold (CAT) method, varies with SNR to minimize the decision error probability in the same way as threshold in CDR based method (γ_d^*) varies with SNR to satisfy the requirement of given probability of detection i.e.$P_d^{th} = 0.9$; whereas the threshold γ_f^* of existing CFAR-based scheme [18] is constant.

Figure 4 is magnified in Fig. 5 to clearly observe the variation of thresholds at low SNR. It is noted that the condition $\gamma_e \in \left[\gamma_f^*, \gamma_d^*\right]$ is satisfied at $\xi \geq -14.5dB$ with sample size of 25,000. At $\xi < -19.5dB$ it is observed that $\gamma_d^* < \gamma_e$, indicates that the optimal threshold does not exist. However, spectrum sensing can still be performed in reliable manner to minimize the error in spectrum sensing and achieve a given value of probability of detection if the number of samples is further increased either through re-sensing or by increasing the sampling frequency.

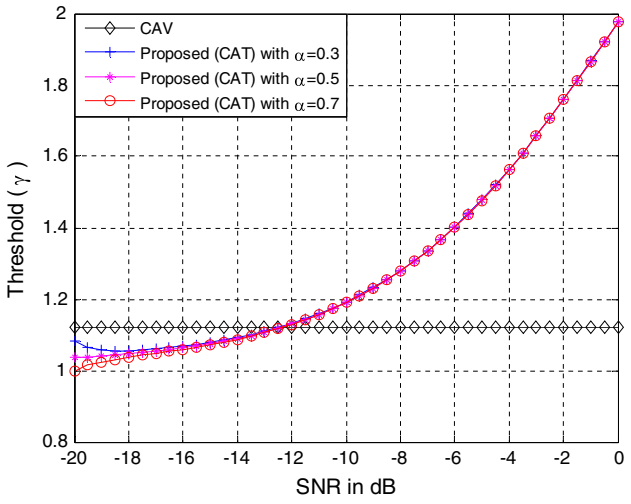


Fig. 2 Threshold value (γ) versus SNR with different spectrum utilization factor of PU

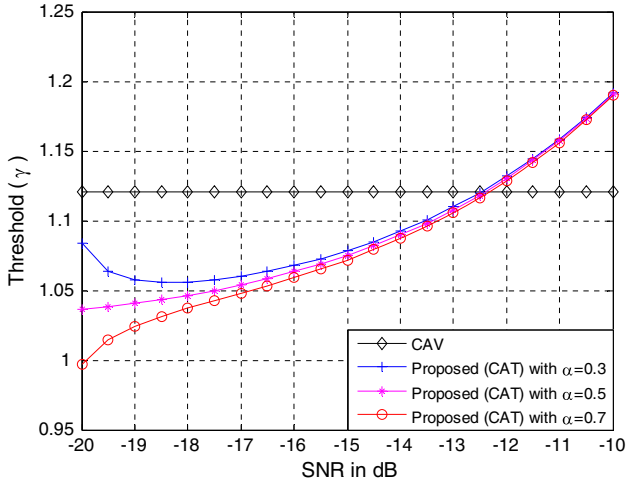


Fig. 3 Magnified version of Fig. 2 at low SNR

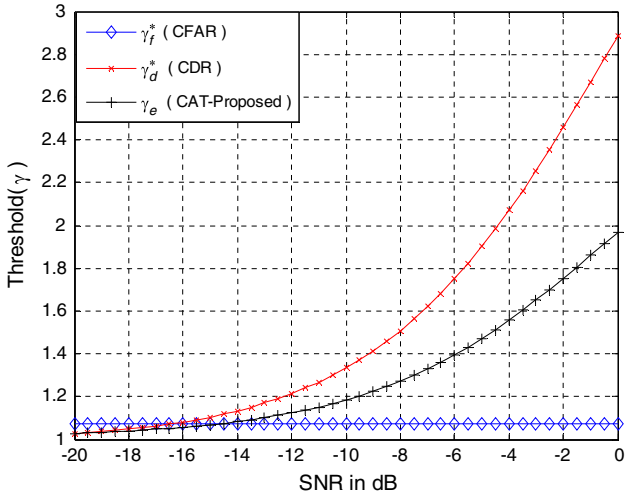


Fig. 4 Threshold value (γ) versus SNR

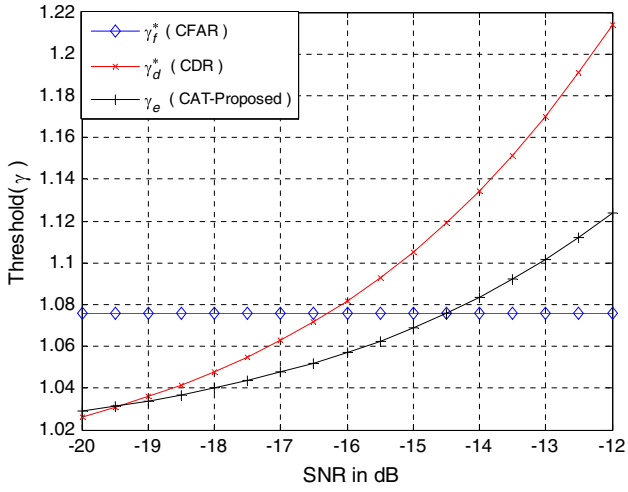


Fig. 5 Magnified version of Fig. 3 at low SNR

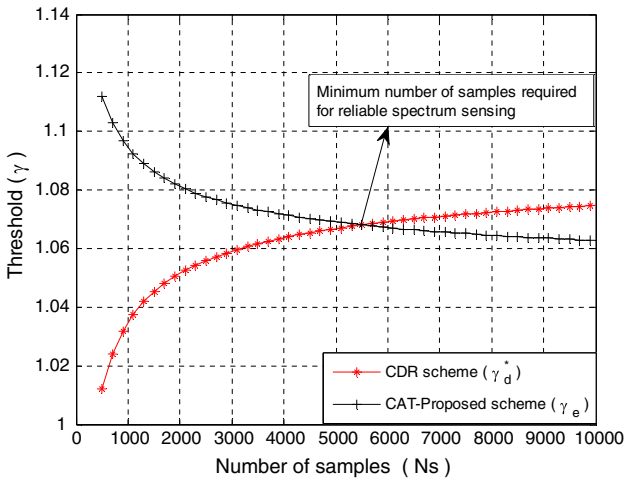


Fig. 6 Threshold value (γ) versus Number of samples

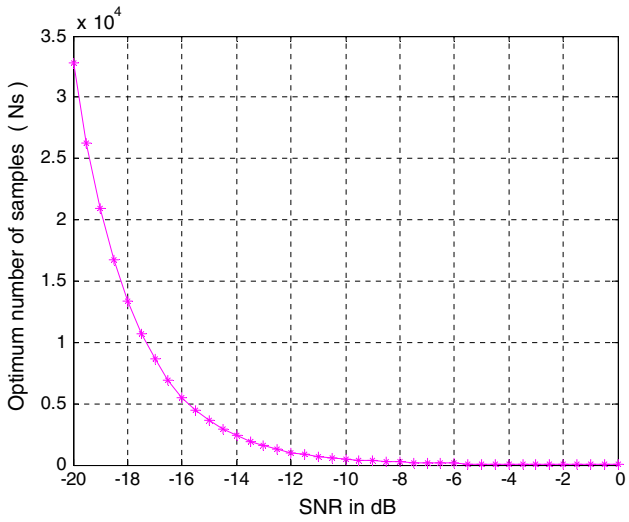


Fig. 7 Required optimum no. of samples for reliable spectrum sensing (N_s) versus Signal to noise ratio

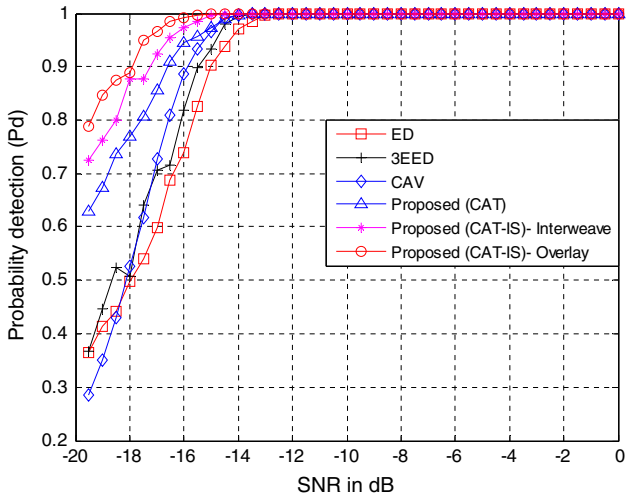


Fig. 8 Probability of detection (P_d) versus SNR

Figure 6 shows the variation of threshold values with number of samples for CDR-based scheme and CAT-based proposed method at $SNR = -16dB$. The minimum number of samples required for reliable spectrum sensing may be obtained from the intersection of curves of CDR and CAT-based thresholds. For examples, the minimum number of samples required at $SNR = -16dB$ is 5500, as shown by Fig. 6. The variation of optimum number of samples required for reliable spectrum sensing with signal to noise ratio is shown in Fig. 7. It can be observed that as the SNR increases fewer samples can provide reliable spectrum sensing.

Figure 8 shows the plot of probability of detection with SNR. From this figure, it is observed that proposed covariance-based adaptive threshold (CAT) scheme performs better than existing CAV method as well as ED-based schemes in absence of any noise uncertainty. The probability of detection is further improved when the threshold is selected in an intelligent manner by using the proposed CAT-IS scheme, especially at low SNR $(\xi < -14dB)$.

For example, the proposed CAT-IS technique achieves a value of P_d at SNR $= -16.5$ dB as 0.954 under the interweave paradigm. Whereas, CAT, CAV, 3EED and ED-based schemes achieve the P_d as 0.911, 0.809, 0.7154 and 0.688, respectively. In order to demonstrate the applicability of the proposed techniques under overlay approach the CAT-IS scheme is also implemented using this approach as a typical example. The CAT-IS technique performs better in overlay approach than interweave as in overlay approach PU shares the information about signal codebook and message with the SU [20–22]. This improves the accuracy of the parameters such as PU's utilization factor and correlation factor of PU's transmitted signal used in the proposed technique. A similar improvement in the performance of CAT scheme may also obtained under overlay approach of spectrum sharing.

In practice the estimated noise power may be different than the actual noise power, therefore some uncertainty is involved in the estimation of noise. If the estimated noise power is assumed to be $\hat{\sigma}_\eta^2 = \beta\sigma_\eta^2$, then the noise uncertainty factor [35, 36] is defined as:

$$B = \max\{10\log_{10}\beta\}dB \tag{33}$$

where, it is assumed that β is uniformly distributed in an interval of $[-B, B]$. In literature, it is assumed that the noise uncertainty factor of 1 dB to 2 dB is normally observed in practice.

For a noise uncertainty of 2 dB, the variation of P_d with SNR is plotted in Fig. 9. From this figure, it can be observed that at high SNR (SNR > -18 dB) the

performance of ED-based method deteriorates most in comparison with other methods under noise uncertainty. However, at low SNR (SNR < -12 dB) the probability of detection remains almost constant ($P_d \approx 0.5$) for ED-based method. In the proposed CAT scheme, the adaptive nature of threshold significantly improves the performance of conventional covariance based (CAV) method especially at low SNR. The proposed CAT and CAT-IS both perform better than CAV and ED schemes at low SNR. When interweave approach is used the better performance of CAT-IS scheme is attributed to appropriate selection of number of samples. When overlay approach is used, the performance of CAT-IS significantly improves in case of noise uncertainty also.

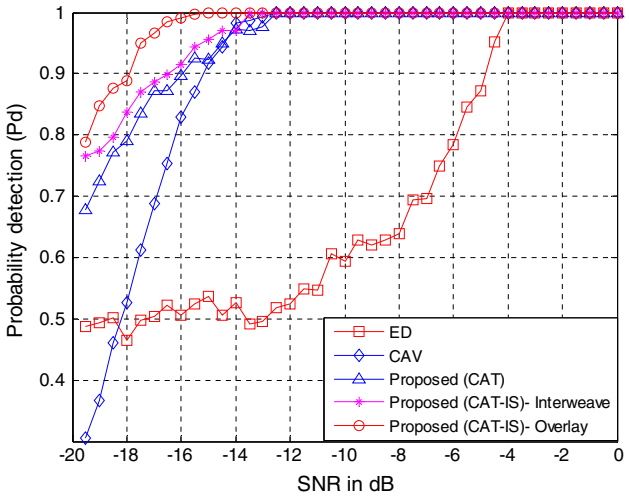


Fig. 9 Probability of detection (P_d) versus SNR, with 2 dB noise uncertainty

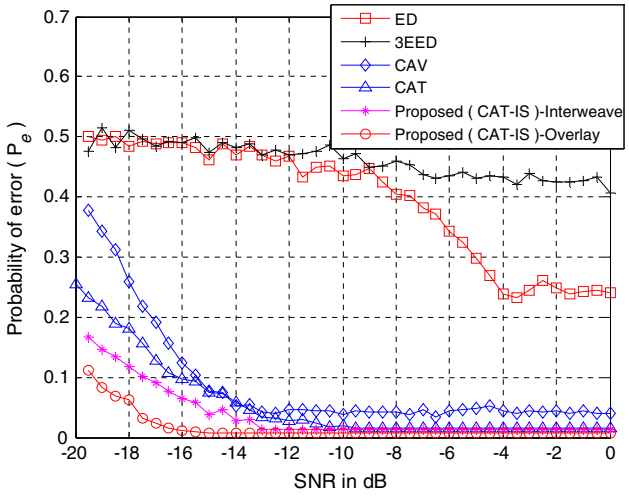


Fig. 10 Probability of error (P_e) versus SNR, with noise uncertainty of 2 dB

Figure 10 shows the variation of probability of error (P_e) with SNR for noise uncertainty of 2 dB. This figure reveals that the performance of ED-based spectrum sensing schemes severely deteriorates under noise uncertainty. The CAT-IS scheme outperforms all other schemes. The best performance is observed when the overlay approach is used due to availability of accurate information about correlation factor and PU's spectrum utilization ratio.

In general, the proposed covariance based adaptive threshold spectrum sensing (CAT) scheme performs significantly better than the existing covariance-based method (CAV) at low SNR even under noise uncertainty, and its performance is much better than ED-based spectrum sensing. This is due to the ability of the proposed algorithm to effectively utilize the information available in terms of spectrum utilization ratio of PU and SNR. Also, the performance of CAT scheme is further improved in CAT-IS method, as threshold is chosen in an intelligent manner by appropriately selecting the number of samples.

As discussed in Sect. 4, the proposed method has similar computational complexity as that of [18]. This fact is also verified by the simulation time required for implementation of these two methods. For example, at $SNR = -16dB$ the simulation time required for proposed scheme is 38% lesser than that required by [18]. Whereas, at $SNR = -19dB$ the proposed method requires 29% more simulation time than that required by the CAV-based spectrum sensing scheme. It is therefore, expected that during the practical implementation also the delay incurred by the proposed method will be similar to that in case of [18]. Thus, it can be concluded that the improvement in the performance is achieved by the proposed method without any increase in computational complexity or delay.

6 Conclusion

In this paper, a scheme for spectrum sensing using an intelligent selection of threshold is presented. The selection process employs three different criteria, namely CFAR, CDR and overall probability of error. First, an adaptive threshold that minimizes the total decision error probability is derived. The closed form expression of the adaptive threshold is based on the statistical covariance matrix of the received signal and varies with spectrum utilization of PU and SNR. The proposed covariance based adaptive threshold (CAT) spectrum sensing scheme significantly improves the performance of existing covariance-based technique and effectively overcomes the problem of noise uncertainty of ED-based spectrum sensing scheme.

Second, to further improve the performance of the spectrum sensing technique, two other thresholds based on CFAR and CDR criteria are also included in the selection

process along with the adaptive threshold. The resulting covariance based adaptive threshold with intelligent selection of threshold (CAT-IS) scheme effectively enhances the performance of CAT technique especially at low SNR and maintains the given probability of detection. The proposed spectrum sensing schemes have computational complexity similar to that of existing CAV method. The applicability of the proposed technique is demonstrated under both interweave and overlay approaches of spectrum sharing. Simulation results have shown that the CAT-IS scheme yields the maximum probability of detection and minimizes probability of overall decision error when compared with existing CAV, conventional ED, 3EED and CAT-based spectrum sensing methods.

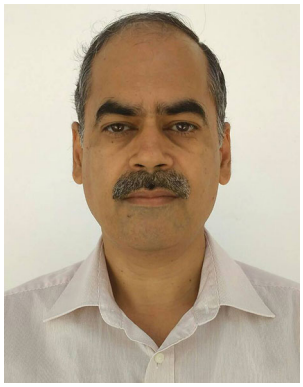
References

1. Federal Communications Commission. (2002) Spectrum policy task force report, FCC 02–155, Nov 2002.
2. FCC. (2003) Facilitating opportunities for flexible efficient and reliable spectrum use employing cognitive radio technologies, notice of proposed rule making and order, in FCC03–322, Dec 2003.
3. Mitola, J., & Maguire, G. Q. (1999). Cognitive radio: Making software radios more personal. *IEEE Personal Communications, 6,* 13–18.
4. Haykin, S. (2005). Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications, 23,* 201–220.
5. IEEE 802.22 Wireless RAN. (2005) Functional requirements for the 802.22 WRAN standard, IEEE 802.22-05/0007r46, Oct 2005.
6. Key, S. M. (1998). *Fundamentals of statistical signal processing: Detection theory* (Vol. 2). Upper saddle: Prentice Hall.
7. Atapattu, S., Tellambura, C., Jiang, H., & Rajatheva, N. (2015). Unified analysis of low-SNR energy detection and threshold selection. *IEEE Transactions on in Vehicular Technology, 64*(11), 5006–5019.
8. Sobron, I., Diniz, P. S. R., Martins, W. A., & Velez, M. (2015). Energy detection technique for adaptive spectrum sensing. *IEEE Transactions on Communications, 63*(3), 617–627.
9. Wang, N., Gao, Y., & Zhang, X. (2013). Adaptive spectrum sensing algorithm under different primary user utilizations. *IEEE Communications Letters, 17*(9), 1838–1841.
10. Vladeanu, C., Nastase, C. V., & Martian, A. (2016). Energy detection algorithm for spectrum sensing using three consecutive sensing events. *IEEE Wireless Communications Letters, 5*(3), 284–287.
11. Kozal, A. S. B., Merabti, M., & Bouhafs, F. (2012) An improved energy detection scheme for cognitive radio networks in low SNR region. *IEEE Symposium on Computers and Communications (ISCC)*, Cappadocia. (pp. 000684–000689).
12. Chen, H.-S., Ga., W. & Daut, D. G. (2007) Signature based spectrum sensing algorithms for IEEE 802.22 WRAN. In *IEEE International Conference Communications (ICC)*.
13. Zhang, X., Gao, F., Chai, R., & Jiang, T. (2015). Matched filter based spectrum sensing when primary user has multiple power levels. *China Communications, 12*(2), 21–31.
14. Oner, M., & Jondral, F. (2004) Cyclostationary-based methods for the extraction of the channel allocation information in a spectrum pooling system. In *Proceedings IEEE Radio and Wireless Conference*, Atlanta, GA. (pp. 279–282).
15. Yang, M., Li, Y., Liu, X., & Tang, W. (2015). Cyclostationary feature detection based spectrum sensing algorithm under complicated electromagnetic environment in cognitive radio networks. *China Communications, 12*(9), 35–44.
16. Zeng, Y., Koh C. L., & Liang, Y. C. (2008) Maximum eigenvalue detection: Theory and application. In *IEEE International Conference on Communications*, 2008. ICC '08. (pp. 4160–4164, 19–23).
17. Charan, C., & Pandey, R. (2016). Eigenvalue based double threshold spectrum sensing under noise uncertainty for cognitive radio. *Optik—International Journal for Light and Electron Optics, 127*(15), 5968–5975.
18. Zeng, Y., & Liang, Y. C. (2009). Spectrum-sensing algorithms for cognitive radio based on statistical covariances. *IEEE Transactions on Vehicular Technology, 58*(4), 1804–1815.
19. Yang, X., Lei, K., Peng, S., & Cao, X. (2011). Blind detection for primary user based on the sample covariance matrix in cognitive radio. *IEEE Communications Letters, 15*(1), 40–42.
20. Jin, M., Li, Y., & Ryu, H. G. (2012). On the performance of covariance based spectrum sensing for cognitive radio. *IEEE Transactions on Signal Processing, 60,* 3670–3682.
21. Yang, X., Lei, K., Peng, S., & Cao, X. (2011). Blind detection for primary user based on the sample covariance matrix in cognitive radio. *IEEE Communications Letters, 15,* 40–42.
22. Jin, M., Guo, Q., Xi, J., Li, Y., Yu, Y., & Huang, D. (2015). Spectrum sensing using weighted covariance matrix in rayleigh fading channels. *IEEE Transactions on Vehicular Technology, 64,* 5137–5148.
23. He., D. (2015) A novel spectrum sensing method in cognitive radio networks based on graph theory. In 2015 *IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, (pp. 1–6).
24. Akhtar, F., Rehmani, M. H., & Reisslein, M. (2016). White space: Definitional perspectives and their role in exploiting spectrum opportunities. *Telecommunications Policy, 40*(4), 319–331.
25. Monemi, M., Rasti, M., & Hossain, E. (2015) Characterizing feasible interference region for underlay cognitive radio networks, In 2015 *IEEE International Conference on Communications (ICC)*, London, (pp. 7603–7608).
26. Monemi, M., Rasti, M., & Hossain, E. (2016). On characterization of feasible interference regions in cognitive radio networks. *IEEE Transactions on Communications, 64*(2), 511–524.
27. Zhang, W., Wang, C. X., Chen, D., & Xiong, H. (2016). Energy–spectral efficiency tradeoff in cognitive radio networks. *IEEE Transactions on Vehicular Technology, 65*(4), 2208–2218.
28. Goldsmith, A., Jafar, S. A., Maric, I., & Srinivasa, S. (2009). Breaking spectrum gridlock with cognitive radios: An information theoretic perspective. *Proceedings of the IEEE, 97*(5), 894–914.
29. Chen, Y., & Oh, H. S. (2016). A survey of measurement-based spectrum occupancy modeling for cognitive radios. *IEEE Communications Surveys and Tutorials, 18*(1), 848–859.
30. Xing, X., Jing, T., Cheng, W., Huo, Y., & Cheng, X. (2013). Spectrum prediction in cognitive radio networks. *IEEE Wireless Communications, 20*(2), 90–96.
31. Saleem, Y., & Rehmani, M. H. (2014). Primary radio user activity models for cognitive radio networks: Survey. *Journal of Network and Computer Applications, 43,* 1–16.
32. Höyhtyä, M., et al. (2016). Spectrum occupancy measurements: A survey and use of interference maps. *IEEE Communications Surveys and Tutorials, 18*(4), 2386–2414.
33. Tumuluru, V., Wang, P., & Niyato, D. (2010) A neural network based spectrum prediction scheme for cognitive radio. In *IEEE International Conference on Communications,* 2012, (pp. 1–5).

34. Ghosh, C., Pagadarai, S., Agrawal, D. P., & Wyglinski, A. M. (2010). A framework for statistical wireless spectrum occupancy modeling. *IEEE Transactions on Wireless Communications, 9*(1), 38–44.
35. Sahai. A., & Cabric, D. (2005) Spectrum sensing fundamental limits and practical challenges. In *Proceedings IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (dySPAN)*, (Baltimore, MD).
36. Tandra, R., & Sahsi, A. (2005) Fundamental limits on detection in low SNR under noise uncertainty. In *Wireless Communications* 2005, (Maui, HI).
37. Liang, Y.-C., Zeng, Y., Peh, E., & Hoang, A. T. (2008). Sensing-throughput tradeoff for cognitive radio networks. *IEEE Transactions on Wireless Communications, 7,* 1326–1337.



Chhagan Charan is working as an Assistant Professor in the Department of Electronics and Communication Engineering, National Institute of Technology, Kurukshetra, India. He received his B. E. (Bachelor of Engineering) degree in Electronics and Communication Engineering from Govt. Engineering College, Bikaner (University of Rajasthan, Jaipur) in 2009 and M. Tech. Degree (Master of Technology) in Microwave and Optical Communication Engineering from Delhi Technological University (DTU), Delhi in 2011. His Research interests include Cognitive Radio, Digital communication, Microwave engg., Wireless sensor networks, and Next generation of communication systems.



Rajoo Pandey is presently working as a Professor in the Department of Electronics and Communication Engineering, National Institute of Technology Kurukshetra, India. He received his Bachelor of Engineering (B.E.) degree in 1989 from Government Engineering College (GEC) Jabalpur, and Master of Technology (M.Tech) from NIT-Kurukshetra (formerly REC Kurukshetra). He received his Ph.D. degree from Indian Institute of Technology, Roorkee in 2001. His research interests include signal processing, communication systems, Digital Signal Processing, Image processing and Cognitive Radio etc.

Available online at www.ijapie.org

INTERNATIONAL JOURNAL OF ADVANCED PRODUCTION AND INDUSTRIAL ENGINEERING

IJAPIE-2017-02-215, Vol 2 (2), 32-41

IJAPIE
Connecting Science & Technology with Management.
A Journal for all Products & Processes.

Assessment of Academic Departments Efficiency Using Data Envelopment Analysis

Sushil Yadav*, Lakshay, Chitransh Joshi, Mandeep Singh, Gurmeet Singh Bhatia , Ranganath M. Singari, Nand Kumar
(Delhi Technological University, Delhi, India)
**Email: sushilyadav5050@gmail.com*

Abstract : Data Envelopment Analysis (DEA) is a relatively new data-oriented approach for evaluating the technical efficiency of a set of peer entities called Decision Making Units (DMUs). A college administration benefits from measurement of relative efficiency of their departments for guiding and supporting inefficient departments. In this study, efficiency measurement of some academic departments within an University. Data Envelopment Analysis (DEA) approach is used for measurement. CCR (Charnes et al. 1978) model of Data Envelopment analysis has been applied for assessing the efficiency of different academic departments using the software "Frontier Analyst". Furthermore, after applying this model for assessment of academic departments efficiency, it was found that Biotechnology Department (Efficiency- 54.8%), Applied Physics Department (Efficiency-89.4%), Civil Department (Efficiency-93.8%), and ECE Department (Efficiency-99.4%) are inefficient Decision Making Units.

Keywords: Data Envelopment Analysis (DEA), technical efficiency, Decision Making Units (DMUs), relative efficiency, CCR model, Frontier Analyst,

1. INTRODUCTION

Higher education is the backbone of development and economic growth in any country. Given that the academic institutions are responsible for the capacity building required for a country's long-term plans, the educational system in particular, is one of the pillars a country depends on to increase its productivity and thus efficiently implement its strategic plans. There is no doubt that an efficient system of institutions of higher education is crucial to providing the necessary profession manpower of scientists, engineers, doctors, and teachers. Therefore, there is a need to assess the efficiency of the educational institutions, and whether the high cost being spent on them is worthwhile. Further, it is necessary to have standards by which all educational institutions could be questioned through the evaluation of efficiency of using resources (Inputs) and achieving the goals (Outputs) for which these resources were spent.

Delhi Technological University (DTU), formerly known as Delhi College of Engineering (DCE), is a premier government university located in New Delhi, India. It is one of the oldest engineering colleges in India and Delhi's first engineering college. It was established in 1941 as Delhi Polytechnic and was under the control of the Government of India. The seeds of Delhi Polytechnic were sowed by the Wood and Abott Committee of 1938. The technical school was created to cater to the demands of Indian industries. At that time, Delhi Polytechnic offered courses in Arts, Architecture, Commerce, Engineering, Applied Science and Textiles. The National Diploma awarded by Delhi Polytechnic was considered an equivalent to the B.E. Degree by the then UPSC. The college has been under the government of the National Capital Territory of Delhi since 1963 and was affiliated with the University of Delhi from 1952 to 2009. In 2009, the college was given state university status, thus changing its name to Delhi Technological University. Till the year 2009, DCE shared its admission procedure and syllabus for various B.E courses with there another branch known as Technology, formerly (DIT), of which were prescribed by Faculty of Technology, University of Delhi. It offers courses towards Bachelor of Technology (B.Tech), Master of Technology (M.Tech), and Doctor of Philosophy (Ph.D.) and contains ten academic departments offering fifteen undergraduate programs namely Computer Engineering, Electronics and Communication, Electrical and Electronics Engineering, Electrical Engineering, Information Technology, Mathematics and Computing Engineering, Software Engineering, Mechanical engineering, Automobile engineering, Civil Engineering, Production and Industrial Engineering, Engineering Physics, Polymer science and Chemical Technology, Biotechnology, Environmental Engineering with a strong emphasis on scientific and technological education and research.

Data Envelopment Analysis (DEA) is a relatively new data-oriented approach for evaluating the technical efficiency of a set of peer entities called Decision Making Units (DMUs). DEA provides a single measure and easily deals with multiple inputs and multiple outputs. Since the DEA technique was first developed, it has been widely applied to industries as diverse as health care (Bhat, Verma, & Reuben,

2001; Jacobs, Smith, & Street, 2006), Banking (Hassan & Sanchez, 2007), and transportation (Pathomsiri, 2006) and many other industries and organizations. Further, DEA-approach has proved especially valuable in cases where we have non-marketed inputs or outputs and/or cannot be derived or agreed upon among different DMUs.
The objective of this study is the "Assessment of Academic Departments Efficiency Using Data Envelopment Analysis"

2. LITERATURE REVIEW

2.1 DATA ENVELOPMENT ANALYSIS

When analyzing system performance, qualitative evaluations such as ``good", ``fair", ``adequate", and ``poor" are vague and difficult to utilize in any meaningful way. As a result, quantitative performance measures are often preferred to such qualitative evaluations. A numerical performance measure might be utilized because the data are readily available, or because it has been used for a long time. Data Envelopment Analysis is one among the several mathematical models to measure flexibility.

DEA involves an alternative principle for extracting information about a population of observations such as those shown in Figure 1. In contrast to parametric approaches whose object is to optimize a single regression plane through the data, DEA optimizes on each individual observation with an objective of calculating a discrete piecewise frontier determined by the set of Pareto-efficient DMUs. Both the parametric and non-parametric (mathematical-programming) approaches use all the information contained in the data. In parametric analysis, the single optimized regression equation is assumed to apply to each DMU. DEA, in contrast, optimizes the performance measure of each DMU. This results in a revealed understanding about each DMU instead of the mythical "average" DMU. In other words, the focus of DEA in on the individual observations as represented by the n optimizations (one for each observation) required in DEA analysis, in contrast to the focus on the averages and estimation of parameters that are associated with single-optimization statistical approaches.

The parametric approach requires the imposition of a specific functional form (e.g., a regression equation, a production function, etc.) relating the independent variables to the dependent variables(s). The functional form selected also requires specific assumptions about the distribution of error terms (e.g., independently and identically normally distributed) and many other restrictions, such as factors earning the value of their marginal product. In contrast, DEA does not require any assumption about the functional form. DEA calculates a maximal performance measure for each DMU relative to all other DMUs in the observed population with the sole requirement that each DMU lie on or below the extremal frontier. Each DMU not on the frontier is scaled against a convex combination of the DMUs on the frontier facet closest to it.

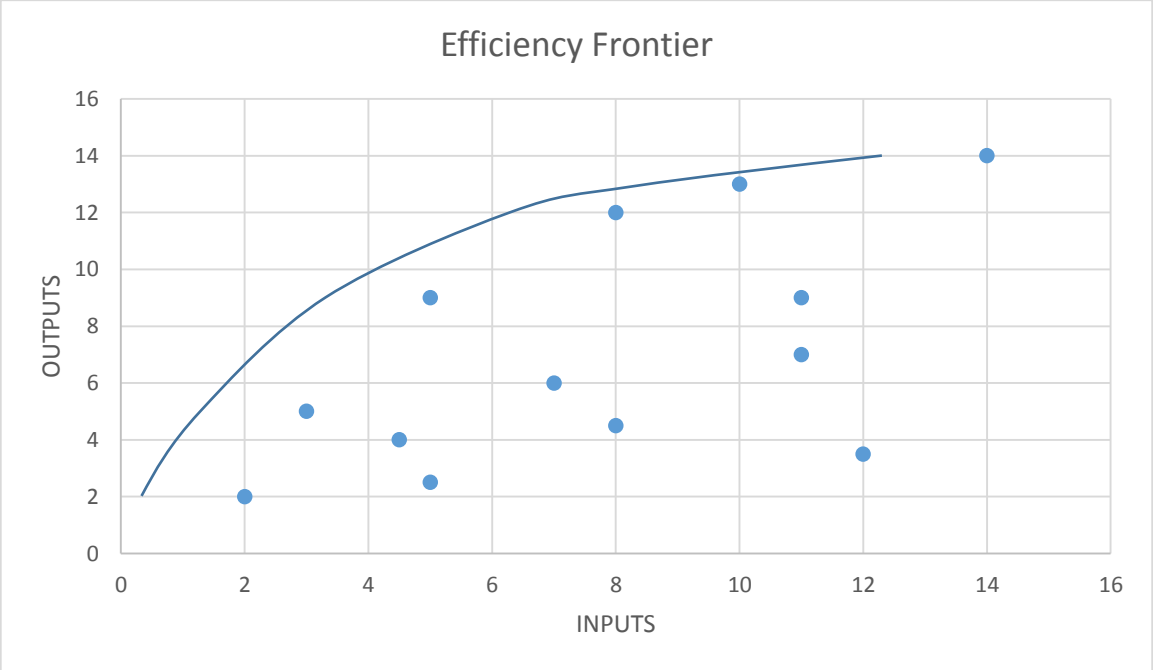


Figure 1

Charnes, Cooper, and Rhodes (1978) extended Farrell's (1957) idea linking the estimation of technical efficiency and production frontiers. Their CCR model generalized the single-output/input ration measure of efficiency for a single DMU in terms of a fractional linear-programming formulation transforming the multiple input/output characterization of each DMU to that of a single "virtual" output and virtual input. The relative technical efficiency of any DMU is calculated by forming the ratio of a weighted sum of outputs to a weighted sum of inputs, where the

weights (multipliers) for both outputs and inputs are to be selected in a manner that calculates the Pareto efficiency measure of each DMU subject to the constraint that no DMU can have relative efficiency score greater than unity.

The solid line in Fig.-1 represents a frontier derived by DEA from data on a population of DMUs, each utilizing different amounts of a single input to produce various amounts of a single output. It is important to know that DEA calculations, because they are generated from actual observed data for each DMU, produce only relative efficiency measures. The relative efficiency of each DMU is calculated in relation to all the other DMUs, using the actual observed values for the outputs and inputs of each DMU. The DEA calculations are designed to maximize the relative efficiency score of each DMU, subject to the condition that the set of weights obtained in this manner for each DMU must also be feasible for all the other DMUs included in the calculation. DEA produces a piecewise empirical extremal production surface (e.g., the solid line in figure 1), which in economic terms represent the revealed best practice production frontier – the maximum output empirically obtainable from any DMU in the observed population, givens its level of inputs.

For each inefficient DMU (one that lies below the frontier), DEA identifies the sources and level of inefficiency for each of the inputs and outputs. The level of inefficiency is determined by comparison to a single referent DMU or a convex combination of other referent DMUs located on the efficiency frontier that utilize the same level of inputs and produce the same or higher level of outputs. This is achieved by requiring solutions to satisfy inequality constraints that can increase some outputs (or decrease some inputs) without worsening the other inputs or outputs. The calculation of potential improvement for each inefficient DMU does not necessarily correspond to the observed performance of any actual DMU making up the piece wise production frontier or to a deterministic projection of an inefficient DMU onto the efficient frontier. The calculated improvements (in each of the inputs and outputs) for inefficient DMUs are indicative of potential improvements obtainable because the projections are based on the revealed best practice performance of "comparable" DMUs that are located on the efficient frontier.

To conclude this section by noting that, initially, DEA relative efficiency solutions were of interest to operational analysts, management scientists, and industrial engineers largely because of three features of this method:

1. Characterization of each DMU by a single summary relative efficiency score;
2. DMU-specific projections for improvements based on observable referent revealed best-practice DMUs; and
3. Obviation by DEA of the alternative and indirect approach of specifying abstract statistical models and making inferences based on residual and parameter coefficient analysis.

The attraction of DEA to traditional Frontier econometricians (Schmidt, 1986, for early concerns) emerged from the new insights obtained in production frontier analysis involving their existence and the variance around them. For example, the BCC model relaxed the constant-returns-to-scale requirement of the original CCR ratio model and made it possible to investigate local returns to scale. Thus it became evident that DEA can complement information about average returns to scale from econometric models with DMU-specific information for most productive scale size (Banker, 1984) and allocative efficiency (Banker and Mainderata, 1988) in contrast to only average most-productive scale and average allocative efficiency obtained from traditional econometric frontier analysis.

2.2 DECISION MAKING UNITS "DMUs"

Data Envelopment Analysis, a non-parametric performance measurement technique, is used for determining the relative efficiency of a set of comparable business called Decision Making Units (DMUs) and it can also be used for benchmarking.

The Delhi Technological University (DTU) has ten departments offering fifteen programs. The ten departments namely Applied Chemistry, Applied Mathematics, Applied Physics, Biotechnology, Civil Engineering, Computer Science & Engineering, Electrical Engineering, Mechanical Engineering, Electronics and Communication Engineering, and Environment Engineering are considered as Decision Making Units (DMUs).

Hence, the research sample includes 10 DMUs offering 15 programs as shown in Table 1.

Table 1 – Decision Making Units

S. No.	Decision Making Units (DMUs)
1.	Applied Chemistry
2.	Applied Mathematics
3.	Applied Physics
4.	Biotechnology
5.	Civil Engineering
6.	Computer Science & Engineering
7.	Electrical Engineering
8.	Mechanical Engineering
9	Electronics and Communication Engineering
10	Environment Engineering

2.4 VARIABLES SELECTION

The efficiency of organizations has been studied by many researchers in different contexts, including research units and university departments. What make efficiency measurement in universities an important issue are legislative requirements, difficulty in determining input and output measures, large number of such measures relative to number of departments, and data collection problems. All these become even more difficult when a college administration desires to compare efficiencies of different academic departments within the college.

A college administration benefits from measurement of relative efficiency of their departments for guiding and supporting inefficient departments. In this study, efficiency measurement of some academic departments within Delhi Technological University. Data Envelopment Analysis (DEA) approach is used for measurement. Necessary criteria and indicators to measure the efficiency of the engineering college departments are described.

Input and output variables: several input-output factors have been considered for this comparative study, they are being enlisted below along with their respective importance and specific roles so as to better appreciate their impact on the overall working efficiency of an academic department.

1. Total Number of Faculty Members (Input 1) - total number of faculty members in any department is one of the most rudimentary inputs by the University in improving the efficiency of the outputs produced. This input has been chosen in the place of staff salaries as suggested by GU¨LSER KO¨KSAL and BURAK NALC¸ACI whose study on Turkish Universities has served as the guiding document. To present a wholesome perspective weighted scores have been given(professor=3pts, associate professor=2pts, assistant prof=1pts)

2. Infrastructure (Input 2) – in order to inculcate the available infrastructure we have included the number of labs that a department has, although several studies have included the area component (Kao and Hung,2006) into their assessment due to empirical constraints we chose the number of labs as the input.

3. Operational Budget (Input 3) – operational budget allotted to different departments can act as a key factor in affecting the output (Kao and Hung,2006) , and is thus included as our last input for this study. Net financial allocations have been directly linked to academic outputs.

4. Total Number of Admissions (Output 1) – total number of admissions is one of the most suited output that an academic department may generate, maximizing it with a given set of inputs can improve efficiency without incurring extra cost. Weighted scores have been included here (Undergraduate= 1pts, Post Graduate= 2pts, PhD candidate= 3pts).

5. Research publications (Output 2) – measuring the efficiency of any academic department shall have research publications as its focus areas and thus this is a desired output.

6. Public Service Activities (Output 3) - the final output is Public Service Activities (PSA)which includes the number of officially documented workshops, conferences, training courses and other extracurricular activities by the teaching staff of each department, that are officially documented at the university (Salah R. Agha, Ibrahim Kuhail ,2011).

2.6 Basic Efficiency Concepts

Efficiency can be simply defined as the ratio of output to input. More output per unit of input reflects relatively greater efficiency. If the greatest possible output per unit of input is achieved, a state of absolute or optimum efficiency has been achieved and it is not possible to become more efficient without new technology or other changes in the production process.

The most common efficiency concept is

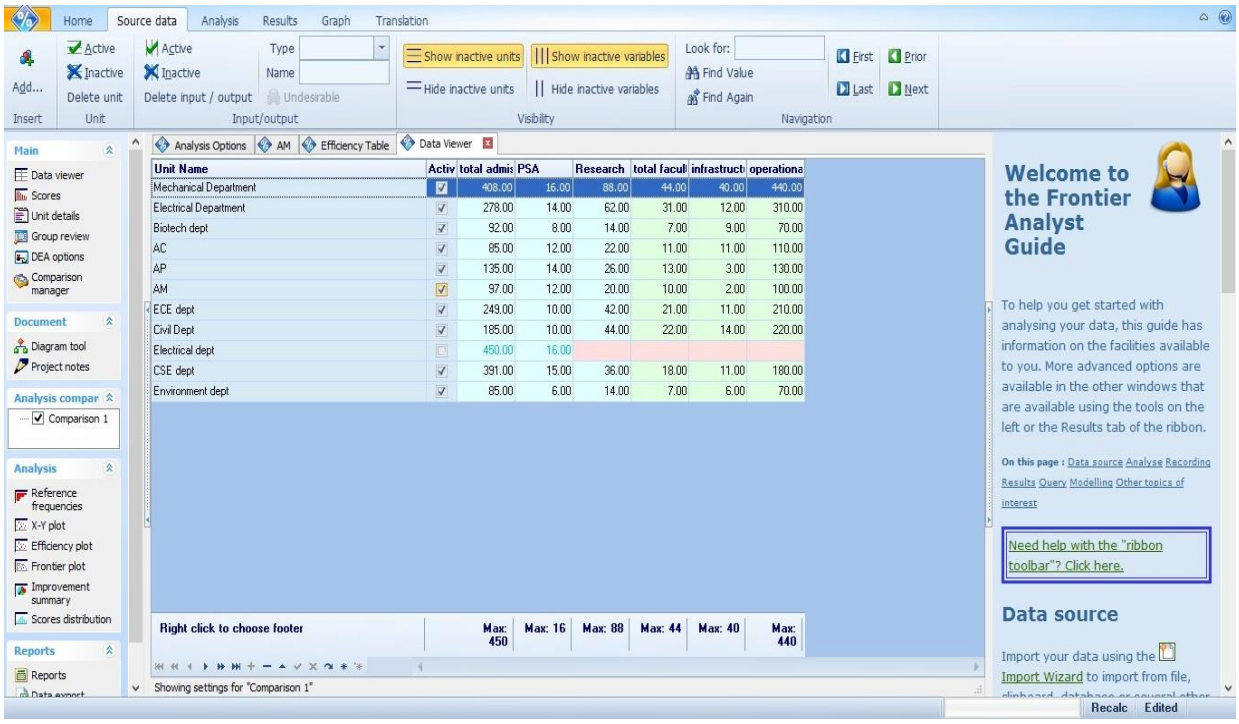
Technical Efficiency: the conversion of physical inputs (such as the services of employees and machines) into outputs relative to best practice. In other words, given current technology, there is no wastage of inputs whatsoever in producing the given quantity of output. An organization operating at best practice is said to be 100% technically efficient. If operating below best practice levels, then the organization's technical efficiency is expressed as a percentage of best practice. Managerial practices and the scale or size of operations affect technical efficiency, which is based on engineering relationships but not on prices and costs.

Allocative Efficiency refers to whether inputs, for a given level of output and set of input prices, are chosen to minimize the cost of production, assuming that the organization being examined is already fully technically efficient. Allocative efficiency is also expressed as a percentage score, with a score of 100% indicating that the organization is using its inputs in the proportions that would minimize costs. An organization that is operating at best practice in engineering terms could still be allocatively inefficient because it is not using inputs in the proportions which minimize its costs, given relative input prices.

Finally, **Cost Efficiency** refers to the combination of technical and allocative efficiency. An organization will only be cost efficient if it is both technically and allocatively efficient. Cost efficiency is calculated as the product of the technical and allocative efficiency scores (expressed as a percentage), so an organization can only achieve a 100%

score in cost efficiency if it has achieved 100% in both technical and allocative efficiency.

Table 2 – Data Collection



2.7 DEA MODEL

We can measure the efficiency of a decision making unit (DMU) simply by comparing its outputs to its inputs. The complex nature of the relation between multiple inputs and multiple outputs involved in the efficiency analysis of DMUs requires utilization of techniques which can handle large number of variables and constraints. For this purpose, Charnes et al. (1978) have proposed DEA. DEA basically discriminates efficient DMUs from inefficient ones. DEA can also point to the improvements each DMU needs to make in each input and output measure. Several studies compare DEA with other methods, and show its importance in measuring relative efficiency of DMUs over regression analysis (Cubbin and Tzanidakis 1998; Seiford and Thrall 1990; Thanassoulis 1993), ratio analysis (Thanassoulis et al. 1996), principle component analysis (Zhu 1998), cluster analysis and discriminant analysis (Stern et al. 1994).

Ali et al. (1995) provide a unified framework for classification of the basic DEA models. DEA models are grouped into CCR (Charnes et al. 1978), BCC (Banker et al. 1984) and Additive Models (Charnes et al. 1985). A DEA model basically identifies among all DMUs the ones that determine an envelopment surface or efficient frontier. The models are classified into two based on the type of envelopment surface: constant returns to scale (such as CCR model) and variable returns to scale (such as BCC model). The CCR model is appropriate when the efficiency of a DMU is not affected by its size (number of personnel, students, and so on). The BCC model, on the other hand, accommodates scale effects in the analysis. DEA models are also grouped into two, depending on the type of proportional movement towards the efficient frontier: input oriented models and output oriented models. Input oriented models address the question "By how much can input quantities be proportionally reduced without changing the output quantities produced?" whereas output oriented models address the question "By how much can output quantities be proportionally expanded without altering the input quantities used?"

The aim of this study is to develop a system to measure the efficiency of college departments and guide the inefficient ones by showing how much they should improve their outputs by using the same amount of inputs or how much they should reduce their inputs by producing the same amount of outputs. Thus, both the input oriented and output oriented model are considered in this case for assessment of the efficiencies of academic departments. Furthermore, the CCR model is preferred, due to less computational effort required by it, and the fact that the number of the departments is very small. The efficiency measure of the output oriented CCR model reflects the ability of a department to obtain maximum output from a given set of inputs and the efficiency measurer of the input oriented CCR model reflects the ability of a department to obtain a given output from minimum set of inputs.

DEA is used to find the relative efficiency of a set of comparable businesses called Decision Making Units (DMU). In DEA efficiency is defined as:

$$\text{Efficiency} = \frac{weighted\ sum\ of\ output}{weighted\ sum\ of\ input}$$

The weight attached to each input or output is not decided beforehand, instead they are chosen to each unit under comparison in its most favorable form. The envelope, or frontier, becomes the surface linking all units whose relative efficiency cannot be exceeded. By definition those units are thus given 100% efficiency, the best possible efficiency for other units in the sample then bring them as close as possible to the envelope. Thus DEA tells us the efficiency of any given DMU relative to best performing units in the sample based on the inputs and outputs.

If there are 'n' DMUs, each with 'm' inputs and 's' outputs, the relative efficiency score of a test DMU 'p' is obtained as follows:

DMU 'p' is obtained as follows:

$$= \text{MAX.} \frac{\sum_{k=1}^{s} V(k).Y(kp)}{\sum_{j=1}^{m} U(j).X(jp)}$$

Such that $\frac{\sum_{k=1}^{s} V(k).Y(ki)}{\sum_{j=1}^{m} U(j).X(ji)} \leq 1$ for all i

(Ref Research Paper 2)

$V(k), U(j) \geq 0$ for all j,k

Where:
K= 1 to's'; j=1 to 'm'; i=1 to 'n'
Y(ki) = amount of output 'k' by DMU 'i'
X(ji) = amount of input 'j' used by DMU 'i'
V(k) = weight given to output 'k'
U(j) = weight given to input 'j'

This can be converted into a linear programming problem for ease of solving, thus, we have

$$\text{MAX} \left(\sum_{k=1}^{s} V(k).Y(ki) \right)$$

Such that $\sum_{j=1}^{m} U(j).X(ji) = 1$

And $\left(\sum_{k=1}^{s} V(k).Y(ki) - \sum_{j=1}^{m} U(j).X(ji) \right) \leq 0$

for all i (Ref Research Paper 2)

The above equations are run 'n' times (one for each DMU) to get the relative efficiency scores of the DMUs. A DMU is said to be efficient if the relative efficiency score is 1, a score <1 insures that it is inefficient. Each DMU selects input and output weights that maximize its relative efficiency.

2.7.1 Benchmarking in DEA

For every efficient DMU, DEA gives a set of efficient DMUs that can be used as benchmark for identifying areas of improvement. The benchmark can be obtained as follows:

Min E sub to –

$$\sum_{i=1}^{n} Q(i)\ Y(ki) \geq Y(kp)$$ for all j

$$\sum_{i=1}^{n} Q(i).X(ki) \leq E.\ X(kp)$$ for all k

$Q(i) \geq 0$ for all i (Ref Research Paper 2)

Where
E= efficiency score
Q(i) = Dual variable

These dual variables can be used to construct Hypothetical Composite Unit (HCU). HCU can be used to measure excess use of inputs or potential increase in outputs. If E= 1, no proof of inefficiency but if E<1 there is efficiency, and thus we can identify the areas of improvement in output and input sides.

DEA is receiving increasing importance as a tool for evaluating and improving the performance of manufacturing and service operations. It has been extensively applied in performance of manufacturing and service operations. DEA approach has the following benefits which make it suitable for its application in flexibility performance measurement and benchmarking.

1. DEA deals with individual cases (Madu and Kuei 1998).
2. It can produces a single measure for each company (Madu and Kuei 1998).
3. It places no restriction on the functional form of the input-output relationship.
4. Able to handle disproportionate multiple inputs and outputs (George and Rangaraj 2008).
5. Does not requiring the decision maker any priory arbitrary weights (George and Rangaraj 2008).
6. It focuses on revealed best-practice frontiers rather than on central tendency properties of empirical data (Madu and Kuei 1998)
7. It can provide an indication of the levels of improvement needed before an inefficient company could be considered efficient (Talluri 2000)

3. ASSESSMENT

3.1 EQUATIONS

Using these Input and output variables equations are formed for ten academic departments, identifying them as DMUs. The standard model of DEA is applied on the following set of equations to compute efficiency.

Testing for inefficiencies in DMU1 by checking for a dummy efficient DMU

Min. E, such that

$W_1 + W_2 + W_3 + W_4 = 1$

$O_{11}xW_1 + O_{12}xW_2 + O_{13}xW_3 + O_{14}xW_4 + O_{15}xW_5 + O_{16}xW_6 + O_{17}xW_7 + O_{18}xW_8 + O_{19}xW_9 + O_{110}xW_{10} >= O_{11}$
$O_{21}xW_1 + O_{22}xW_2 + O_{23}xW_3 + O_{24}xW_4 + O_{25}xW_5 + O_{26}xW_6 + O_{27}xW_7 + O_{28}xW_8 + O_{29}xW_9 + O_{210}xW_{10} >= O_{21}$
$O_{31}xW_1 + O_{32}xW_2 + O_{33}xW_3 + O_{34}xW_4 + O_{35}xW_5 + O_{36}xW_6 + O_{37}xW_7 + O_{38}xW_8 + O_{39}xW_9 + O_{310}xW_{10} >= O_{31}$

$I_{11}xW_1 + I_{12}xW_2 + I_{13}xW_3 + I_{14}xW_4 + I_{15}xW_5 + I_{16}xW_6 + I_{17}xW_7 + I_{18}xW_8 + I_{19}xW_9 + I_{110}xW_{10} <= I_{11}$ (E)
$I_{21}xW_1 + I_{22}xW_2 + I_{23}xW_3 + I_{24}xW_4 + I_{25}xW_5 + I_{26}xW_6 + I_{27}xW_7 + I_{28}xW_8 + I_{29}xW_9 + I_{210}xW_{10} <= I_{21}$ (E)
$I_{31}xW_1 + I_{32}xW_2 + I_{33}xW_3 + I_{34}xW_4 + I_{35}xW_5 + I_{36}xW_6 + I_{37}xW_7 + I_{38}xW_8 + I_{39}xW_9 + I_{310}xW_{10} <= I_{31}$ (E)

Testing for inefficiencies in other DMUs on similar lines we get the following results, as given below. Assessing them gives us the scope of improvement and identifies efficient DMUs, this helps in adopting techniques of operations management and productivity enhancement from the efficient DMUs to the relatively inefficient ones.

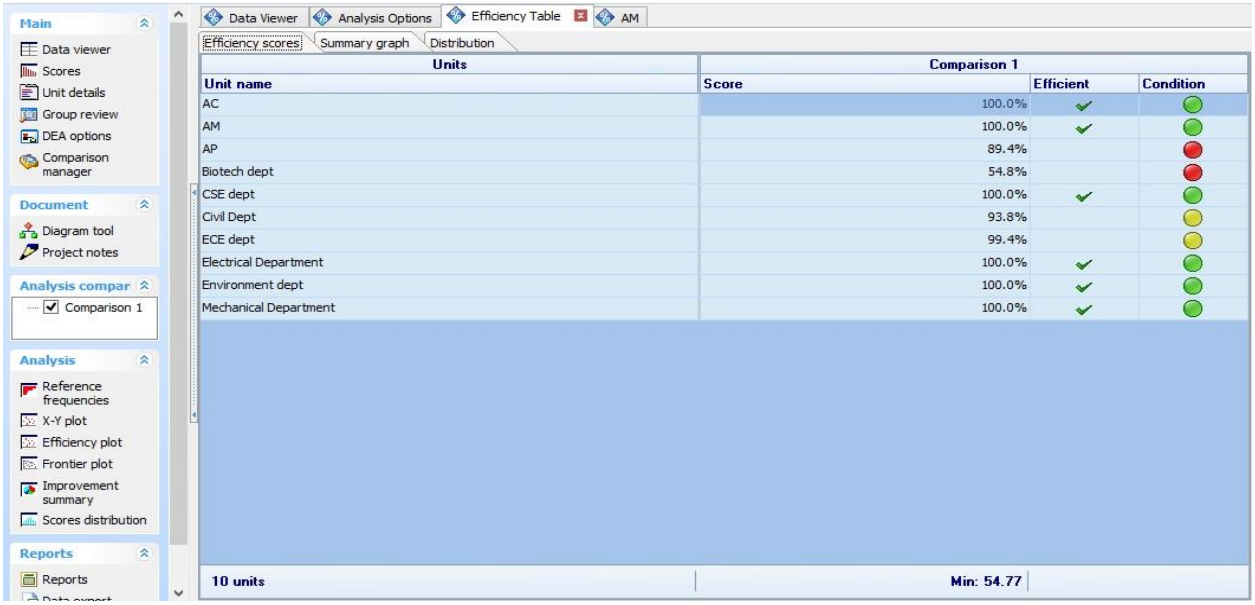


Figure 2 – Efficiency Score

As can be seen that the relative efficiencies have been color coded to identify the ones that need focus immediately, wherein red color denotes urgent attention with vast scope of improvement, yellow indicates a mild scope of improvement, and green indicates that the DMU lies on the efficiency frontier.

3.2 EFFICIENCY PLOTS

As Biotech dept. is the one with the least score it is analyzed first. Here it can be suggested that with an input optimization perspective either a reduction in the inputs is needed or a corresponding increase in the outputs is required to justify existing inputs.

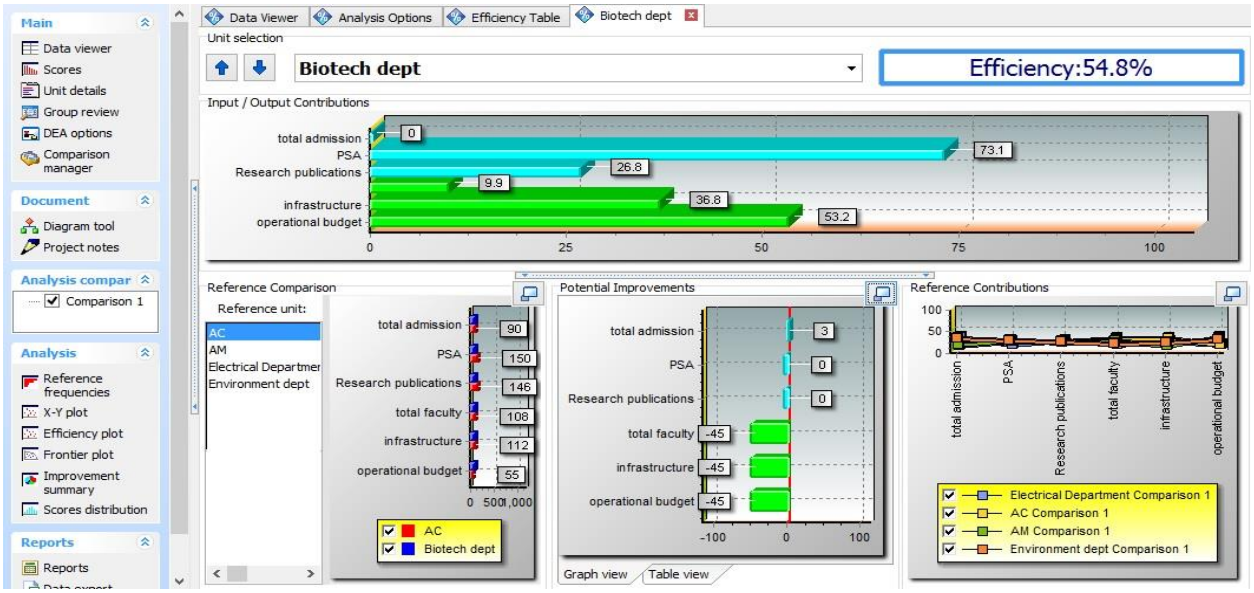


Figure 3 – Efficiency Plot of Biotechnology Department

Now looking at the efficiency record of Applied Physics department we can suggest a lay of in terms of faculty as the output levels don't justify the high levels of existing faculty. Suitable reductions in strength of the faculty and operational budget have been shown here.

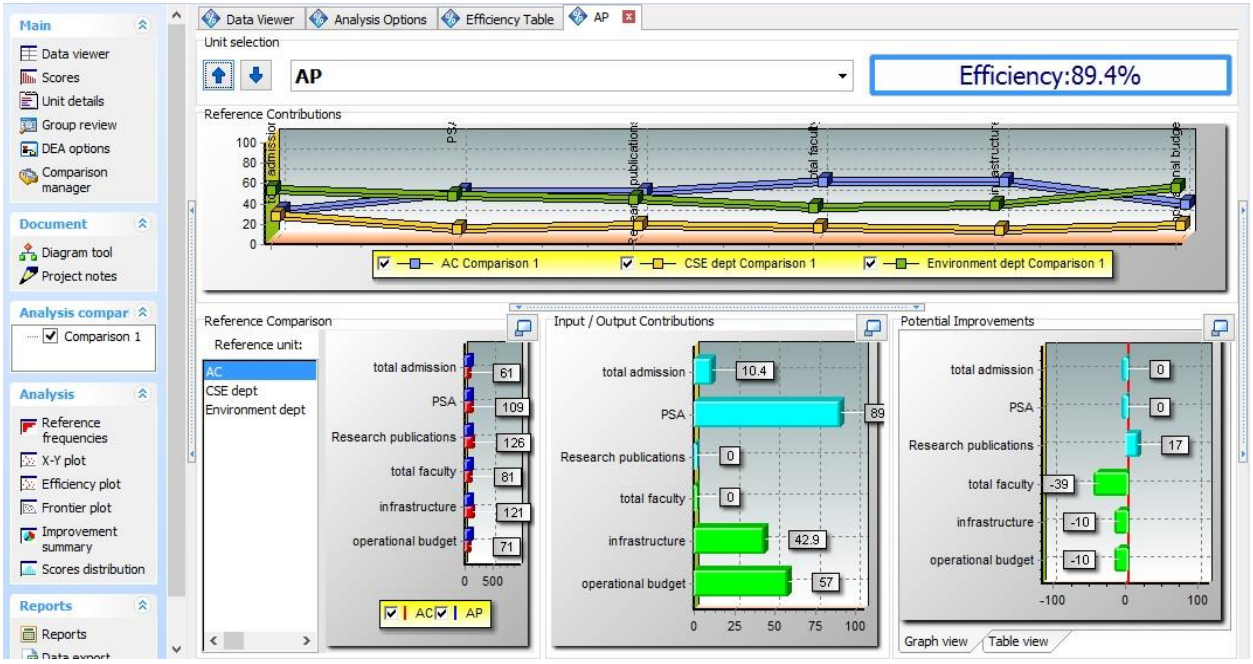


Figure 4 – Efficiency Plot of Applied Physics Department

Further taking an example from the yellow coded DMUs, we look at the Civil Department, here the efficiency relative to the others is fairly high (93.8%), but still a revision of the input given can be considered given the slight fall in efficiency.

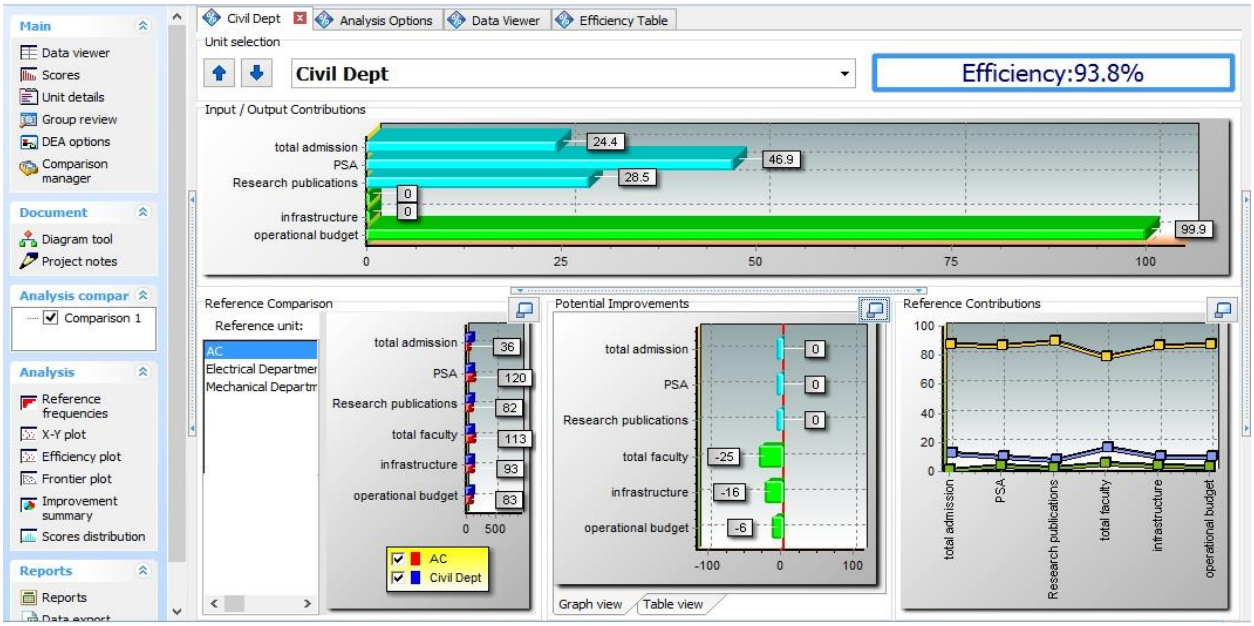


Figure 5 – Efficiency Plot of Civil Department

The total potential improvements for the Biotech Department have been specifically highlighted here in order to lay special emphasis on it due to its score. The best direction to start with in improvements would be to reduce the faculty and operational budget.

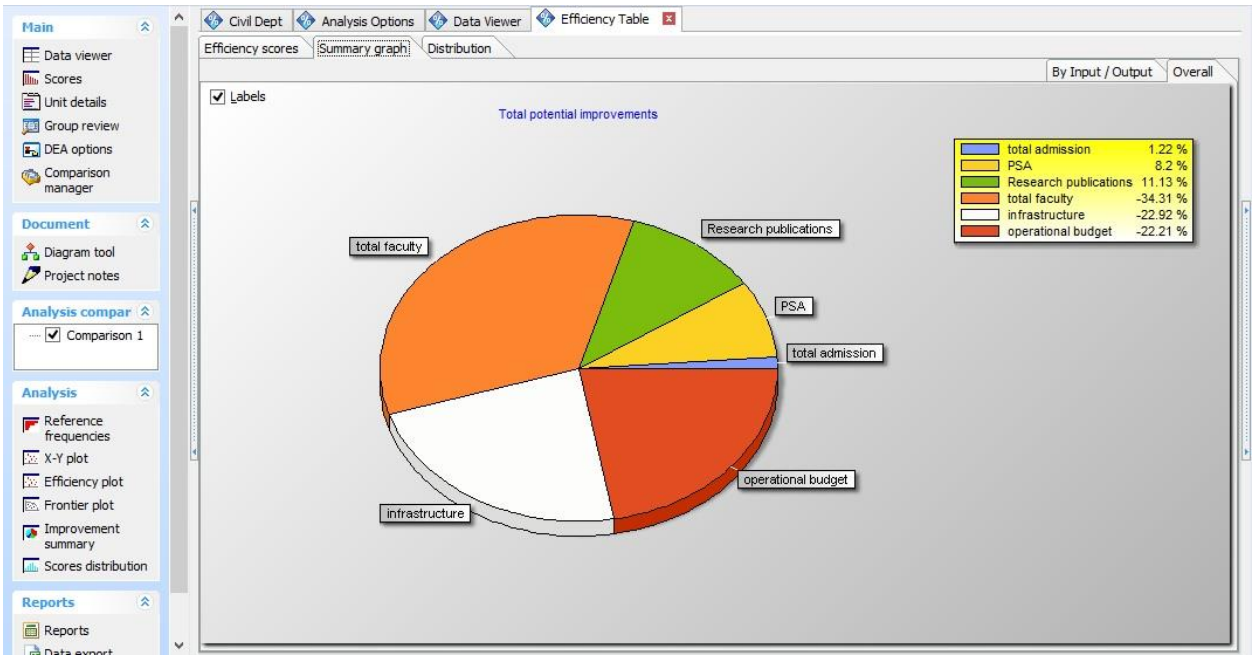


Figure 6 – Total Potential Improvements

4. CONCLUSIONS

During this project, we have gone through the research literature about Data Envelopment Analysis. CCR (Charnes et al. 1978) model of Data Envelopment analysis has been applied for assessing the efficiency of different academic departments using the software "Frontier Analyst". Furthermore, after applying this model for assessment of academic departments efficiency, we found that Biotechnology Department (Efficiency- 54.8%), Applied Physics Department (Efficiency-89.4%), Civil Department (Efficiency-93.8%), and ECE Department (Efficiency-99.4%) are inefficient Decision Making Units. Various steps should be taken to identify the bottlenecks in these cases and suggestion to rectify the same and thus improve the overall efficiency.

REFERENCES

1. Ying-Ming Wang, Kwai-Sang Chi: A neutral DEA model for cross-efficiency evaluation and its extension, in Expert Systems with Applications, An International Journal (2010),ref 3666–3675
2. Salah R. Agha, Ibrahim Kuhail, Nader Abdelnabi, Mahmoud Salem, Ahmed Ghanim: Assessment of academic departments efficiency using data envelopment analysis. Journal of Industrial Engineering and Management (2011), p301-325.
3. GU¨LSER KO¨KSAL & BURAK NALC¸ACI," The relative efficiency of departments at a Turkish engineering college: A data envelopment analysis ", International Journal of Operations & Production Management (2006) 51: 173–189
4. Rita Markovits-Somogyi," ranking efficient and inefficient decision making Units in data envelopment analysis ", Industrial Management & Data Systems, (2011)pp. 446 – 456
5. Abbott, M., & Doucouliagos, C. The Efficiency of Australian Universities: A Data Envelopment Analysis. Economics of Education Review, 22(1),(2003), 89-97..
6. Lopes, A., Lanzer, E. Data Envelopment Analysis-DEA And Fuzzy Sets to Assess The Performance of Academic Departments: A Study at Federal University Of Santa Catrina-UFSC. Pesquisa Operacional, 22(2), (2002), 217-230.
7. Moreno, A. & Tadepalli, R.. Assessing Academic Department Efficiency a Public University. Managerial and Decision Economics, 23(7), 385-397.8. Janusz Bucki Yvon Pesqueux, (2000),"Flexible workshop: about the concept of flexibility", International Journal of Agile Management Systems, Vol. 2 Iss 1, (2002) pp. 62 – 70
9. Tom Davis (1993) ,"Effective Supply Chain Management",Sloan Management Review, Summer 1993 edition.
10. Nand Kumar, Ranganath.M.S, Chandrashekhar, "Sustainability Analysis of Banking Sector in India using DEA", in National Conference on Inclusive & Sustainable Growth - IMT Nagpur
11. www.emeraldinsight.com
12. Banxia Frontier Analyst 4 Demo version.

Nuclear Instruments and Methods in Physics Research B xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research B

journal homepage: www.elsevier.com/locate/nimb

Lifetime quenching due to surface wake field

Gaurav Sharma [a,b], Nitin K. Puri [a,*], T. Nandi [b]

[a] Dept. of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India
[b] Inter University Accelerator Centre, New Delhi 110067, India

ARTICLE INFO

Article history:
Received 15 December 2016
Received in revised form 14 May 2017
Accepted 29 May 2017
Available online xxxx

Keywords:
Surface wake field
Life time measurement
Beam foil spectroscopy

ABSTRACT

We report the measurement of surface wakefield on the lifetime of He-like Ti ($1s2s\ ^3S_1$) during the time-of-flight measurements in the beam-foil experiment. Quenching in the lifetime of $1s2s\ ^3S_1$ state from 26.6 ns to 67.11 ps has been observed due to blending with short-lived $1s2p\ ^3P_1^o$ states in the presence of wakefield present near the target surface. The measured and theoretical lifetime of unperturbed He-like Ti ($1s2s\ ^3S_1$) and ($1s2p\ ^3P_1^o$) has been used to calculate the blending parameter ε (3.74×10^{-2}). The surface wakefield for 110 MeV Ti^{+8} passing through 80 μg/cm^2 thick carbon foil is found to be 9.15×10^7 V/cm. The measured stark mixing parameter can be used to determine the surface wakefield as well as the wake stopping power of an ion through a solid target.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Beam-foil time-of-flight [1–3] is a well known experimental technique to determine the lifetime of a state through photon intensity decay curve of moving ions. The beam-foil excitation is not a state selective process [2]; therefore, most of the spectra depart from the single exponential behaviour and show multi-exponential or power law [4,5]. Apart from the non-selective excitation nature of the beam foil; cascade from the higher levels makes the lifetime measurement more difficult and sometimes may show an unusual sudden enhancement in the intensity [6,7]. However, these are not the only problem with the measurement of the lifetime of a state using beam foil spectroscopy. The electric field generated at the exit-surface of the foil perturb the nearby states and therefore deprive the measurements from obtaining the true lifetime of the state [8–10]. The origin of this electric field is explained by the loss of energy by fast electrons to the bulk and surface of solid targets through surface Plasmon theory [11] and the surface stopping power is expected due to the wake potential [12,13] originated from the surface Plasmon. The energy loss field in the bulk of the foil is sufficiently high to blend the fine-structure sub-levels completely; however, the contribution from the surface is quite less in comparison to the bulk field [14]. This surface wakefield (SWF) induced blending may lead to a partial admixture of nearby states [10]. This blending can be utilized to measure the SWF, the stopping range of ions and the energy losses at the surface of the target. In this paper, we have discussed the results of time-of-flight measurement using the beam-foil collisional spectroscopy. The lifetime of He-like Ti $1s2s\ ^3S_1$ is measured. The present observations show an observable change in the lifetime of $1s2s\ ^3S_1$ state of He-like Ti ions from nanosecond to picosecond. This change in the lifetime is attributed to the partial blending of $1s2s\ ^3S_1$ and $1s2p\ ^3P_1^o$ states. The blending of states is measured in terms of blending parameter ε [15,5] and then the SWF intensity involved in the blending is calculated using the blending parameter value.

2. Experimental setup

The experiment was performed at IUAC, New Delhi, India using the 15 unit doubled particle accelerator having palletron geometry [16]. An ion beam of 110 MeV Ti^{8+} beam was passed through 80 μg/cm^2 thick C target to excite or strip off the electrons to create highly charged and excited ions. The thickness of the target was selected such that the charge equilibrium is achieved. The thickness required to charge equilibrium was estimated using Monte-Carlo based ETACHA code [17]. The self-supported carbon (C) foils were prepared using the electron gun evaporation technique and annealed at 900 K in the N_2 environment to obtain better uniformity and mechanical stability. The C target on Al holder was mounted on a target movement assembly (lifetime setup) having a translational accuracy better than 0.1 μm. An Al target holder, instead of steel, was selected to avoid any spurious peaks from the target holder in the spectra (as the X-rays of aluminium (Al) are far separated from Ti for our detection system). Two low energy Germanium X-ray detectors were positioned outside the vacuum

* Corresponding author.
E-mail address: nitinpuri2002@yahoo.co.in (N.K. Puri).

http://dx.doi.org/10.1016/j.nimb.2017.05.063
0168-583X/© 2017 Elsevier B.V. All rights reserved.

Please cite this article in press as: G. Sharma et al., Lifetime quenching due to surface wake field, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.063

at an angle of ±90° to the beam direction; to keep the measurements reliable (shown in Fig. 1). The X-rays were collected from the beam-foil source through a 6 μm thick Mylar window. Collimators were installed in the system to make the measurements precise. A Faraday cup having electron suppressor was used to normalize the X-ray spectra during lifetime measurements. The X-ray absorption technique was employed in the experiment for improving the resolution and reduce pileup effects present at the high energy end of the spectra. A detailed explanation of this method with some experimental results is explained elsewhere [18], where fractional absorption is used to enhance the intensity of low-intense transitions. In the present experiment, a 10.2 ± 0.20 μm thick Aluminum (Al) filter was used as an absorber in front of the X-ray detectors. The filters follow the Beer Lambert's law to attenuate the intensities, so it always reduces a fixed percentage of incident photons of particular energy. Therefore the nature of a decay curve does not change after employing the filters, and hence the lifetime information. The lifetime of the 1s2s $^3S_1$ state was also confirmed without using any filter. In both the cases, the lifetime was measured to be quenched to the picoseconds. The optical zero of the system was ascertained by the maximum intensity in the decay plots to avoid any uncertainty in the zero position. The vacuum in the chamber was stable in the range of 10^{-7} Torr throughout the experiment.

3. Discussion

A prompt X-ray spectrum observed with the above mentioned experimental environment is shown in the Fig. 2. The first two peaks in the Fig. 2 are the K_α and K_β transitions of He-like projectile ions in 110 MeV Ti ion and C atoms collisions. The detector used in the experiment does not have sufficient resolution to resolve H, He and other species. The transitions may also have some contribution from the H-like and Li-like ions also. To check the contribution from H-like Ti theoretically; the charge state fraction is required. The charge state fraction is estimated using the latest available theoretical code, ETACHA4 [19] for the H-like ions and is found to be only 0.06% of the all Ti charge states. This little contribution from H-like ions is omitted. Also, the fitting of first peak does not converge for He- and Li-like Ti transitions. Therefore keeping this in view, the peak is confirmed as He-like Ti.

The transition at 4.73 keV persists for the few mm of the delay during the lifetime measurements length which in turn equivalent

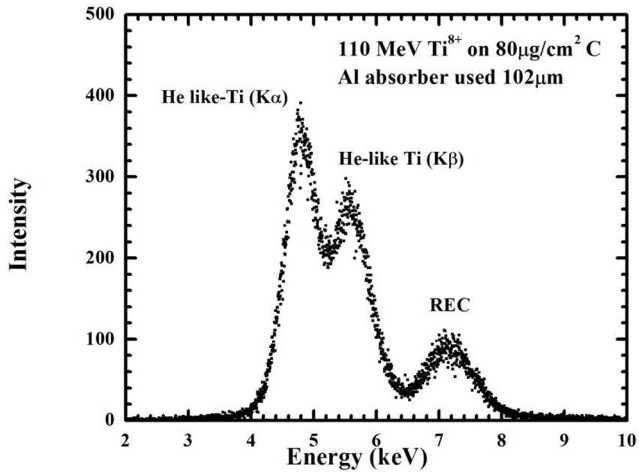


Fig. 2. The X-ray spectrum of 110 MeV Ti on 80 μg/cm^2 C target. First peak is a combination of 1s2s $^3S_1$ and 1s2p $^3P_1$.Third peaks is due to REC photon decaying to ground state in Ti ions. Second peak requires more experimental confirmation before assignment.

to the tens of ps. The transition probability of the He-like K_α (4.73 keV) transition is $2.40 \times 10^{+14}$ s^{-1} [20] and lifetime of femtosecond order [21,22]. Therefore the contribution from slowly decaying levels is also expected in the 4.73 keV transition. The lifetime of these transitions are measured by delaying beam-foil source from the detector which is 67.11 ± 8.07 ps. A typical X-ray spectrum is shown in Fig. 2.

The energy of the second peak (5.50 keV) matches with the K_β (1s3p $^3P^o_1 \rightarrow$ 1s^2 $^1S_0$) and (1s3s $^3S_1$ – 1s^2 $^1S_0$) transition in He-like Ti [20]. The ion energy in the experiment was near Coulomb barrier, so, X-rays from projectile-like ions are also expected. The nuclear product Cr is one of the candidates for nuclear product forms by the alpha transfer reaction [23,24]. The energy of the second peak is also matched with the K_α transition from He-like Cr [20], however, the detector resolution was not good enough to resolve. So, there is confusion in the assignment of the second peak (5.50 keV) in the spectrum due to competition between He-like Ti and He-like Cr ions (nuclear product). Secondly, the intensity ratio for K_α and K_β should be 15 [25], but the value does not match with the present spectrum analysis. Therefore the second peak is not discussed in the present manuscript as it requires high-

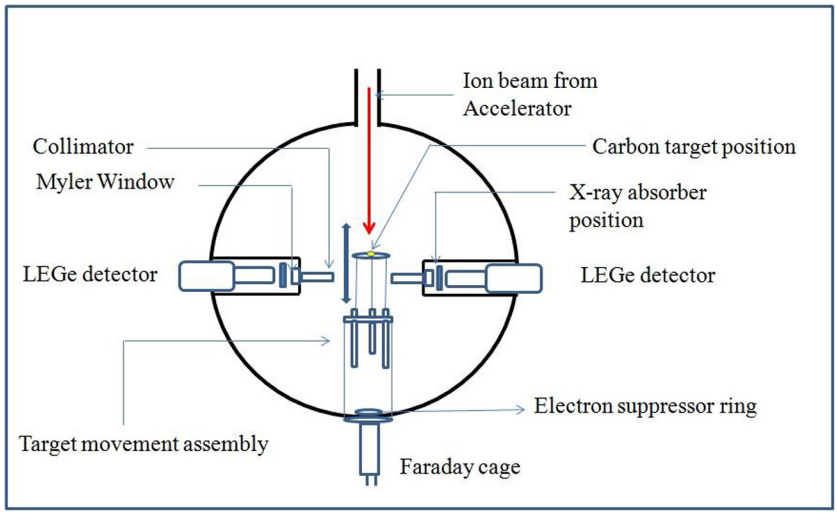


Fig. 1. Experimental setup used for the measurement of Lifetime using delay method. Two X-ray detectors were symmetrically placed to avoid any error from detector. Position accuracy of system was better than 1 μm.

Please cite this article in press as: G. Sharma et al., Lifetime quenching due to surface wake field, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.063

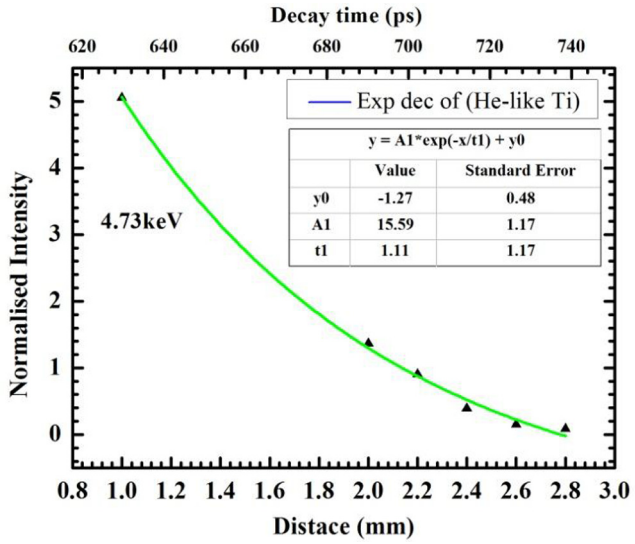


Fig. 3. The decay curves for the projectile and projectile like ions for the 110 MeV Ti ions on 80 μg/cm^2 C. The curves have been confirmed with both the detectors.

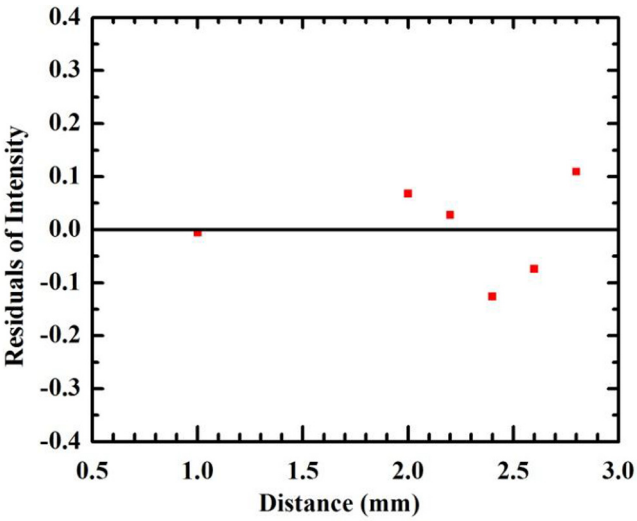


Fig. 4. The residuals of the single exponential fitting shown in Fig. 3 are plotted.

resolution spectroscopy and more theoretical supports. The third peak is assigned to radiative electron capture (REC) photon decay. The REC peak in the spectrum is much broader than the other two. This broadening is due to the distribution of charge states inside the carbon target. Here each state leads to the radiative capture of electron and thereafter REC photons of different energies are emitted out [26].

The lifetime of He-like Ti (1s2s $^3S_1$) having energy 4.73 keV has been measured [20]. The intensity decay curve is fitted to a single exponentially decaying function as shown in Fig. 3. Residuals of the exponential fit vary within ±0.1 value as shown in Fig. 4. Using the decay constant (1.11 ± 0.17 mm from Fig. 3) and the beam velocity (21.04 mm/ns) for 110 MeV $^{48}Ti^{8+}$ beam the lifetime of the (1s2s $^3S_1$) state is calculated as 67.11 ± 8.07 ps while the expected value is 26.6 ns [21,22]. Such a large change in the lifetime and hence in the transition probability of the state requires perturbation from some other nearby short-lived level. The electric field present at the exit C target surface can stark mix the two nearby levels easily. The closely separated 1s2s $^3S_1$ and 1s2p $^3P^o_1$ states satisfy the criteria of opposite parity [27] for stark mixing leading to change the transition rate. The strength of field and separation between the states decide the extent of blending. If the field is high enough, it can blend the two interacting states completely; and individual identity of the states is lost. This upper limit of the field can be calculated by $\Delta E = \frac{3Fn(n-1)}{Z}$ [15] where F is the field required to lose the individual identity of the states completely. If a field comparable to F can be achieved by the field then a sufficient change can be observed in the lifetime [10]. Now, the 1s2p $^3P^o_1$ and 1s2s $^3S_1$ energy levels of the He like Ti are separated by 22.6 eV [20]. Such closely spaced levels separated by few tens of eV and having opposite parity can stark mix under the influence of electric field. This stark mixing leads to change the transition probability of both the states. Therefore the departure from the expected lifetime is due to the stark mixing of the closely spaced 1s2p $^3P^o_1$ and 1s2s $^3S_1$ states in the case of He-like transitions (peak 1). The 1s2p $^3P_1$ to ground state is an E1 transition having a short lifetime of the order of picoseconds or shorter. On the other hand, $^3S_1$ to the ground state is the forbidden transition having longer lifetimes of the order of the nanosecond. The interaction between these two short and long lived levels must be considered for calculating the interaction parameter. Therefore, for a perturbed system, the relationship between the observed and theoretical rates should be

$$\psi'_{1s2s\,3S_1} = \psi_{1s2s\,3S_1} + \varepsilon\psi_{1s2p\,3P^o_1},$$

$$\omega'_{1s2s\,3S_1} = \omega_{1s2s\,3S_1} + (\varepsilon)^2\omega_{1s2p\,3P^o_1}$$

Where $\omega' and \omega$ are the measured and theoretical transition rates, respectively. The parameter ε is the blending parameter. The values of transition probability (ω) is used from NIST database [20]. The formula is used previously for H-like cases [10,28] and here we have assumed that the formula is valid until the interaction is between two states and hence valid for He-like case also. The value of ε is found to be 3.74×10^{-2} for blending of He-like Ti (1s2p $^3P^o_1$, 1s2s $^3S_1$) states. The electric field strength involved for the change in the lifetime or blending of the corresponding two closely lying states are calculated from $\varepsilon = \frac{E'}{475*Z^5}$ [15,5], Where E′ is the field concerned in the blending of two states and ε is the blending parameter. The field required for 100% blending (F) is calculated 2.06×10^9 V/cm, by putting ε value equals to 0.7071. The value of SWF is estimated to be 9.15×10^7 V/cm and shown is Table 1; which is less than the value for complete blending ($\sim 10^9$ V/cm).

4. Conclusion

The measurement of SWF induced stark mixing of 1s2s $^3S_1$ and 1s2p $^3P^o_1$ states in He-like Ti is observed using the time-of-flight

Table 1
The expected and measured lifetime of the corresponding ions are given. The blending parameter 'ε' and strength of electric field involved 'E' for the blending is calculated for the ions. A is the transition rate and ΔE is the energy difference between the two blending states.

Ion species	E (keV)	Expected Lifetime (ns)	Measured lifetime (ps)	ε	E (V/cm)	Blending state	ΔE (eV)	A (s^{-1}) [20]
He like Ti 1s2s $^3S_1$	4.73	26.60	67.11 ± 8.07	3.74E−02	9.15E + 07	1s2p $^3P^0_1$	22.60	1s2s $^3S_1$-3.76E+07 1s2p $^3P^0_1$-1.06E+13

Please cite this article in press as: G. Sharma et al., Lifetime quenching due to surface wake field, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.063

technique during the passage of 110 MeV Ti^{8+} ions through 80 μg/cm^2 thick C target. This stark mixing leads to quenching in the lifetime of 1s2s $^3P_1$ level from of 26.6 ns to 67.11 ± 8.07 ps. The measured and the theoretical lifetimes of these state has been used to determine the blending parameter ε (3.74×10^{-2}), after that it is utilized to calculate the SWF intensity. The calculated SWF ($E = 9.15 \times 10^7$ V/cm) is found to be two order smaller than the field ($F \sim 10^9$ V/cm) required to blend the states completely. This blending parameter can also be utilized to determine the wake stopping power of an ion through solid target [10].

References

[1] R. Wehlitz, P.N. Jurani, Phys. Rev. A 74 (2006) 42721.
[2] H.G. Berry, Annu. Rev. Nucl. Part. Sci. 32 (1982) 1.
[3] T. Nandi, P. Marketos, P. Joshi, R.P. Singh, C.P. Safvan, P. Verma, A. Mandal, A. Roy, R.K. Bhowmik, Phys. Rev. A 66 (2002) 52510.
[4] C. Can, R.J. Maurer, B. Bandong, R.L. Watson, Phys. Rev. A 35 (1987) 3244.
[5] R.W. Dunford, H.G. Berry, D.A. Church, M. Hass, C.J. Liu, M.L.A. Raphaelian, B.J. Zabransky, L.J. Curtis, A.E. Livingston, Phys. Rev. A 48 (1993) 2729.
[6] T. Nandi, J. Phys. B: At. Mol. Opt. Phys. 42 (2009) 125201.
[7] N.N. Nedeljković, S.M.D. Galijaš, M.A. Mirković, J. Quant. Spectrosc. Radiat. Transf. 134 (2014) 46.
[8] T. Nandi, Open J. Microphys. 2011 (2011) 53.
[9] T. Nandi, K. Haris, G. Singh, P. Kumar, R. Kumar, S. Saini, S. Khan, A. Jhingan, P. Verma, A. Tauheed, D. Mehta, H.G. Berry, Phys. Rev. Lett. 110 (2013) 163203.
[10] T. Nandi, B.P. Mohanty, J. Phys. B: At. Mol. Opt. Phys. 42 (2009) 225402.
[11] R.A. Baragiola, C.A. Dukes, P. Riccardi, Nucl. Instr. Meth. B 182 (2001) 73.
[12] F.J.G. de Abajo, P.M. Etchenique, Phys. Rev. B 45 (1992) 8771.
[13] A. Koyama, Phys. Rev. Lett. 65 (1990) 3156.
[14] F.J. Garcia de Abajo, P.M. Echinique, Phys. Rev. A 46 (1992) 2663.
[15] Hans A. Bethe, Edwin E. Salpeter, Quantum Mechanics of One- and Two-Electron Atoms, 1977.
[16] D. Kanjilal, S. Chopra, M.M. Narayanan, I.S. Iyer, V. Jha, R. Joshi, S.K. Datta, Nucl. Instr. Meth. A 328 (1993) 97.
[17] J.P. Rozet, C. Stephan, D. Vemhet, Nucl. Instr. Meth. B 107 (1996) 67.
[18] G. Sharma, D. Swami, B. Kumar, N.K. Puri, T. Nandi, Arxiv 1512.08399 (2015).
[19] E. Lamour, P.D. Fainstein, M. Galassi, C. Prigent, C.A. Ramirez, R.D. Rivarola, J.P. Rozet, M. Trassinelli, D. Vernhet, Phys. Rev. A 92 (2015) 42703.
[20] NIST Atomic Spectra Database, http://www.nist.gov/pml/data/asd.cfm.
[21] J. Hata, I.P. Grant, J. Phys. B: At. Mol. Opt. Phys. 14 (1981) 2111.
[22] J. Hata, I.P. Grant, J. Phys. B: At. Mol. Opt. Phys. 16 (1983) 507.
[23] H.J. Karwowski, S.E. Vigdor, W.W. Jacobs, S. Kailas, P.P. Singh, F. Soga, T.G. Throwe, T.E. Ward, D.L. Wark, J. Wiggins, Phys. Rev. C 25 (1982) 1355.
[24] A.P. Mishra, T. Nandi, B.N. Jagatap, J. Quant. Spectrosc. Radiat. Transf. 118 (2013) 70.
[25] A. Thompson, D. Attwood, E. Gubikson, and Malcolm howells, Sect. 1.2 X-RAY Emiss. ENERGIES 45, 2001.
[26] B.S. Datz, B.R. Appleton, C.D. Moak, Atomic Collisions in Solids, second ed., Springer Science, n.d.
[27] R.D. Cowan, Theory At. Struct. Spectra, 1981, pp. 14.13.
[28] Y. Kumar, A.P. Mishra, T. Nandi, J. Quant. Spectrosc. Radiat. Transf. 148 (2014) 22.

Please cite this article in press as: G. Sharma et al., Lifetime quenching due to surface wake field, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.063

Iran Polym J
DOI 10.1007/s13726-017-0537-x

ORIGINAL RESEARCH

Microencapsulation of reactive amine by interfacially engineered epoxy microcapsules for smart applications

Manorama Tripathi[1,2] · Devendra Kumar[2] · Prasun Kumar Roy[1]

Received: 19 December 2016 / Accepted: 29 May 2017
© Iran Polymer and Petrochemical Institute 2017

Abstract The most common approach for incorporation of extrinsic self-healing functionality relies on introducing healant-loaded micro-containers in the polymeric formulation. In this context, a healing system based on encapsulated epoxy resin and amine hardener appears to be one of the most economically viable solutions, in view of the chemical as well as mechanical compatibility with the matrix. Encapsulation of epoxy resins has been extensively studied while the high reactivity of the amine hardener renders its encapsulation rather difficult and has been attempted with only modest success. The purpose of the present work is to adopt an interfacial polymerization approach for the preparation of epoxy microcapsules encapsulating a reactive amine hardener (triethylene tetramine). The effects of experimental parameters, including reaction temperature, stirring speed and epoxy/amine concentration ratio on the microcapsule formation were investigated. A polymeric surfactant was used to stabilize the suspension to modulate the particle size distribution of the resultant microcapsules. The highest encapsulation efficiencies were obtained when the reaction medium was maintained at 70 °C under stirring (600 rpm) at epoxy/amine ratio of 10/3.2. The microcapsule dimensions and core content could be tailored, following this encapsulation approach of interfacial polymerisation. Under optimal conditions, spherical microcapsules with 100% yield and 12% core content were obtained.

Keywords Microencapsulation · Interfacial polymerization · Microstructure · Surfactant · Core content

Introduction

In view of the exceptional mechanical properties offered by polymeric composite materials, their use in structural applications is increasing steadily in areas mandating high specific strength, particularly automotives, wind turbine blades, etc. Epoxy resin is the preferred choice as the matrix material for preparation of fiber-reinforced composites; however, the cross-linked resin is highly sensitive to crack initiation. The cracks initiate at nano-scale, which subsequently enlarge to micro-scale and finally to macro-scale until total failure ensues. To increase the serviceability and reliability of the polymeric structural composites, development of self-healing polymers as an area received intense attention by the scientific community over the last decade.

The economic viability of healing in the structure of the composites requires cost-effective chemical approaches that are robust, efficient and stable. In this context, three primary approaches are in vogue towards introducing self-healing functionality to polymers including extrinsic, intrinsic and vascular [1]. The most common approach for incorporation of extrinsic self-healing functionality relies on introducing healant-loaded micro-containers in the polymeric formulation [2, 3]. Alternately, healing agents can also be encapsulated in 3D microvascular channels [4]. Many

Electronic supplementary material The online version of this article (doi:10.1007/s13726-017-0537-x) contains supplementary material, which is available to authorized users.

✉ Prasun Kumar Roy
pk_roy2000@yahoo.com

[1] Centre for Fire, Explosive and Environment Safety, DRDO, Timarpur, Delhi 110054, India

[2] Department of Applied Chemistry and Polymer Technology, Delhi Technological University, Delhi 110042, India

Published online: 07 June 2017

research works have been focused on the binary systems, in which two components meet in the damage zone, react and trigger the healing response. Several microcapsule-based systems have been reported such as DCPD/Grubbs catalyst [5, 6], ENB/Grubbs catalyst [7, 8] DCPD/tungsten hexachloride [9], PDMS/dimethyldineodecanoate tin catalyst [10], isocyanate/thiol [11], PDMS/Pt catalyst [12], epoxy/mercaptan [13], epoxy/boron trifluoride diethyletherate ($(C_2H_5)_2O{\cdot}BF_3$) [14] and unsaturated polyester/AIBN [15]. Approaches relying on the latent functionality have also attracted enormous research interest [16].

Another combination based on encapsulated epoxy resin and amine hardener appeared to be an extremely attractive healing system, which has been reported to be compatible with epoxy matrix both chemically and mechanically. Studies on the preparation of microcapsules encapsulating epoxy resin have been extensively reported [17, 18]. However, in view of the high reactivity of the amine hardener, its encapsulation in aqueous medium is rather difficult and has been attempted with only modest success. In this context, successful encapsulation of liquid amine in polyurea shell using a reverse emulsion technique has been previously reported [19]. The potential of interfacial polymerization of amine with isocyanate to form polyurea microcapsules have also been investigated [20, 21]. An alternative explored route involved synthesis of hollow poly(urea-formaldehyde) microballoons which was subsequently filled with amine [22]. Recently, etched glass balloons have also been used as microcontainers for reactive amines [23]. Literature survey clearly highlights that development of microcapsules encapsulating amines with demonstrable healing effect is an area of great interest to industry as well as academia.

In our previous papers, we have demonstrated the applicability of interfacial polymerisation route towards encapsulation of healants like benzoxazine and epoxy resins in epoxy shells [24, 25]. We considered it of interest to adopt this technique towards encapsulation of water-soluble amines, a process far more challenging than the ones previously reported. In this study, we extended the scope of our technique by reporting the preparation of mechanically robust epoxy micro-containers encapsulating triethylenetetramine (TETA) hardener and explored its potential toward healing. A polymeric surfactant was employed to stabilize the amine–oil suspension for modulating the particle size of microcapsules. Encapsulation of the amine was achieved by its interfacial reaction with epoxy to form a cross-linked shell, leaving behind excess amine unreacted within the core, which can subsequently serve as an epoxy curing agent for smart applications.

Experimental

Materials

Aliphatic epoxy resin (Ciba Geigy, Araldite CY 230; epoxy equivalent 200 eq g^{-1} Basel, Switzerland) and TETA-based hardener (Ciba Geigy, HY 951; amine content 24 eq kg^{-1} Basel, Switzerland) were used as received. Liquid paraffin (Heavy) (Central Drug House, Mumbai, India, kinematic viscosity 300 cSt) and polymeric surfactant, Hypermer A70 (Croda, Snaith, UK) were used without further purification. Double distilled water was used throughout the course of experimental work.

Microencapsulation of amine

The immiscibility of triethylene tetramine in liquid paraffin was substantiated by capturing the optical images of the suspensions prepared by stirring vigorously for 5 min (1000 rpm) using a computer interfaced microscope (magnification 40×). Subsequently, cross-linked epoxy microcapsules filled with amine were prepared by interfacial polymerization of TETA hardener with epoxy resin in non-stoichiometric ratio, while keeping them suspended in liquid paraffin oil. The reaction was carried out under varying temperatures (50–75 °C), stirring speeds (400–600 rpm) and varying resin/hardener ratios (100/12–100/3.2). Based on the stoichiometry, a epoxy: amine ratio of 100:12 was required, as established by calculations presented in the supplementary information file.

For the purpose of encapsulation, the amine hardener (5 g) was initially suspended in paraffin (100 mL) under ambient condition while stirring at 600 rpm. The temperature of the medium was slowly increased to the desired temperature, followed by introduction of predetermined amount of epoxy resin through a hypodermic syringe. The reaction was allowed to proceed for extended periods, followed by cooling to room temperature and the resultant microcapsules were filtered, washed extensively and air-dried. The ratio of mass of the microcapsules obtained to the total mass of reactants (epoxy resin and hardener) was calculated to quantify the microcapsule yield. The encapsulated amine within the free flowing microcapsules could be extracted by refluxing in water for 12 h. The difference in the mass due to extraction, as estimated gravimetrically, was used as a quantification of core content.

Characterization

The surface morphology of amine-encapsulated microcapsules was studied using a scanning electron microscope

(EVO MA15, Zeiss, Germany) under an acceleration voltage of 1 kV. The samples were mounted on the aluminum stubs and sputter-coated with gold and palladium (10 nm) using a sputter coater (Quorum SC7620, Oberkochen, Germany) operating at 10–12 mA for 120 s. Average microcapsule size and standard deviations were determined by measuring the diameters of about 100 microcapsules using ImageJ software. Optical images of suspended microdroplets were captured using optical microscope (MOTIC B3-223PL, Kowloon, Hong Kong).

Structural characterization of microcapsules was performed by recording FTIR spectra in the wavelength range of 4000–600 cm^{-1} using a Thermo scientific FTIR (NICOLET 8700, Madison, USA) analyser with an attenuated total reflectance (ATR) crystal accessory. A Perkin Elmer Diamond STG-DTA (Shelton, Connecticut, USA) apparatus was used to investigate the thermal degradation behavior of the amine-filled microcapsules under N_2 atmosphere in the temperature range 50–800 °C at a heating rate of 10 °C/min. A sample mass of 5.0 ± 0.5 mg was used for each experiment.

Calorimetric studies were performed on a differential scanning calorimeter (TA instruments Q 20, Hertfordshire, UK). For dynamic DSC scans, samples (10 ± 2 mg) were sealed in aluminum pans, and heated from 0 to 300 °C at 10 °C/min. All experiments were performed under N_2 purge (50 mL/min) to minimize oxidative degradation of the sample. Prior to the experiments, the instrument was calibrated for temperature and enthalpy using standard indium and zinc. Thermal equilibrium was regained within 1 min after sample insertion, and the exothermic reaction was considered to be completed when the recorder signal leveled off to the baseline. The total area under the exothermic curve was determined to quantify the heat of curing.

Results and discussion

Amine-filled microspheres were prepared by the reaction of a cycloaliphatic epoxy with triethylene tetraamine (TETA) using inverse suspension polymerization technique utilizing the principles of interfacial engineering. The effects of operating parameters, particularly stirring speed, temperature and epoxy/hardener ratio on the microcapsule dimensions, yield and core content were studied.

Microencapsulation of amine-based hardener

The most common encapsulation technique relies on the formation of a polymeric shell around the liquid healant suspended in a dispersion medium. For example, in the extensively studied poly(urea-formaldehyde) microencapsulation technique, the shell wall-forming monomers, namely urea and formaldehyde react in the aqueous phase to form a low molecular weight prepolymer. As the molecular weight increases, the soluble prepolymer deposits at the healant–water interface which leads to the formation of a smooth inner surface of the shell. The cross-linking process further continues to form a stable shell wall around the healant [26]. The rough exterior of the shell resulted from the precipitation of higher molecular weight prepolymer formed in the aqueous solution, which subsequently aggregate and deposit on the capsule surface.

Unfortunately, this encapsulation methodology is not successful for amine encapsulation, in view of the solubility of the amine in aqueous medium, which does not permit the formation of a stable separate phase. For encapsulation of these water-soluble substances, inverse emulsion polymerization (water in oil) is the norm [27], which inspired us to explore the possibility of interfacial engineering approach. In the present study, the water-soluble amine was dispersed in paraffin medium followed by the controlled introduction of reactive epoxy in non-stoichiometric amounts, lesser than the amount required for complete curing. Microcapsules were formed by the preferential reaction of the epoxy with TETA at the paraffin–amine interface. A schematic of the entire procedure is presented in Fig. 1.

The use of paraffin as the dispersion medium has been well reported [27], although not yet for amine-based systems. Theoretical predictions were made to assess the possibility of amine–paraffin phase separation. The solubility parameter of epoxy, TETA and paraffin was calculated based on the group contribution models of Hoy [28]. The structures of the molecules along with the detailed calculations are presented in the supplementary information file. The difference in the solubility parameters of TETA and paraffin, i.e., 10.8 and 8.25 $cal^{1/2}$ $cm^{3/2}$ mL^{-1}, respectively, suggests the existence of TETA as phase



Fig. 1 Schematic of the encapsulation process

separated droplets under agitation in paraffin medium. To verify the same, optical images of TETA–paraffin suspension were captured (Fig. 2), which clearly indicate the immiscibility of TETA in paraffin.

It is to be noted that stabilization of the inverse emulsion pose a challenge in microcapsule synthesis [19]. In view of the same, we were not surprised to observe the instability of the amine–paraffin dispersion where the coalescence of the amine droplets was evidenced (Fig. 2a). Different surfactants were explored towards their ability of stabilizing the obtained dispersion, for example sodium dodecyl sulphate (SDS), Pluronic 61 (PEO–PPO–PEO), polyvinyl alcohol (PVA) and Hypermer A70, of which the last three surfactants are polymeric. The optical images of the dispersions obtained using surfactant are presented in the supplementary information file (Fig. S1). It can be seen that Hypermer A70 was the most effective surfactant in stabilizing the amine droplets. The uniformity in the size distribution is presented in Fig. 2b.

Upon introduction of epoxy resin to the amine–paraffin suspension, ring opening of the oxirane functionality initiates at the oil–amine interface resulting in the formation of prepolymer around the amine droplets. The reaction then proceeds with the diffusion of both the monomers (epoxy and hardener) towards the interface, up to the complete consumption of the limiting reagent.

It is to be noted that the microcapsule dimensions are strongly dependant on the experimental parameters. Extensive studies were performed to arrive at the optimal conditions, especially in terms of reaction time, temperature and stirring speed. Prior to encapsulation experiments, the curing behavior of the epoxy resin in the presence of TETA was investigated using non-isothermal calorimetric studies. The effect of increase in heating rate viz. 2.5, 5, 10 and 15 °C min^{-1} on the curing behavior of the neat epoxy is presented in Fig. 3a.

Characteristic thermal parameters, i.e., ΔH_{cure}, T_{onset} and T_{peak} are summarized in the supplementary information file (Table S1). As expected, increase in the heating rate (β) led to a systematic shift in the curing behavior towards higher temperatures. The Kissinger–Akahira–Sunose method [29] was applied for determination of the activation energies (E_α) from the plot of $\ln\left(\frac{\beta_i}{T^2_{\alpha,i}}\right)$ against $1/T_\alpha$, as follows:

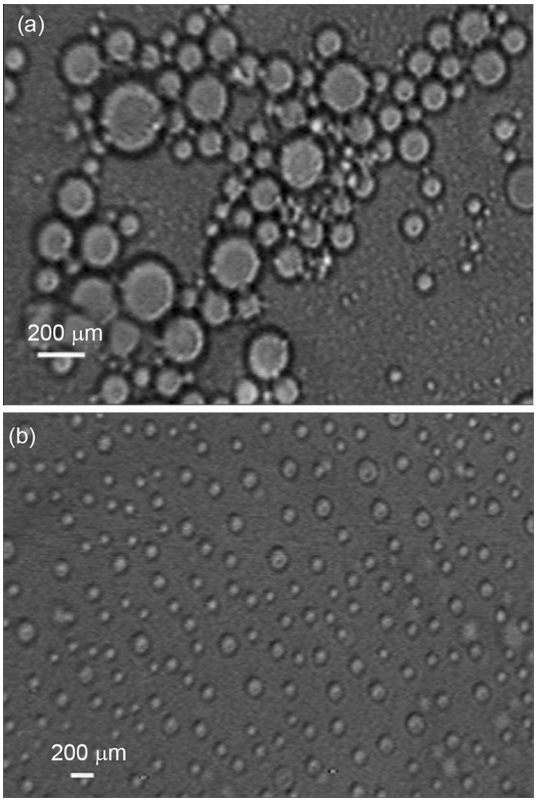


Fig. 2 Optical images of TETA–paraffin suspension in: **a** absence of surfactant and **b** presence of surfactant

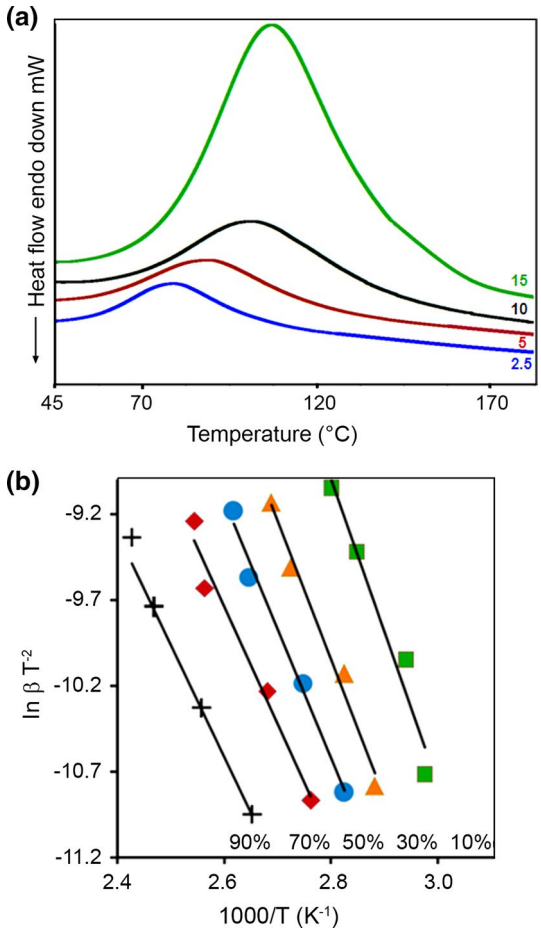


Fig. 3 **a** Effect of heating rate on the curing profile of epoxy and **b** isoconversional plots at different conversions using Kissinger–Akahira–Sunose method

$$\ln\left(\frac{\beta_i}{T_{\alpha,i}^2}\right) = \text{Const} - \frac{E_\alpha}{\text{RT}_{\alpha,i}}, \quad (1)$$

where β is the heating rate, T_α is the temperature associated with a particular conversion (α) at the corresponding heating rate, E_α is the activation energy at a particular degree of conversion (α) and R is the gas constant. The linear plots at few representative conversions are also presented in Fig. 3b, and the variation of activation energy with conversion is presented in the supplementary information file (Fig. S2) [30]. Our studies clearly reveal that the exothermic curing process can be initiated under ambient conditions; however, it would require impractical durations, in view of which encapsulation studies were performed over the temperature range 50–75 °C.

The SEM micrographs of the prepared microspheres using stoichiometric amounts of epoxy and amine (epoxy/amine = 10/1.2), obtained at different temperatures are presented in Fig. 4. It seems that at lower temperatures (T < 65 °C) only partial curing takes place, even after extended periods (~36 h) which was evidenced by the formation of agglomerated microcapsules. As the temperature of the reaction media was raised to 70 °C, the extent of curing increased, leading to the formation of free flowing microcapsules within 20 h. Increase in the temperature further led to significant vaporization of the amine, which in turn led to further imbalance in stoichiometry over extended time periods. To support this, a set of isothermal thermogravimetric experiments were performed on TETA, where the mass loss was studied as a function of time at different temperatures ranging from 50 to 90 °C. Few representative isothermal TG traces are presented in the supplementary information file (Fig. S3). As expected, mass loss increased with increase in temperature. Over extended durations, especially at higher temperatures, the vaporization of amine was expected to lead to an imbalance of stoichiometry, which in turn resulted in partial curing and agglomeration of the microcapsules. Therefore, detailed studies on the effect of other operating parameters, i.e., stirring speed and epoxy/hardener ratio were performed while maintaining the reaction mixture at 70 °C.

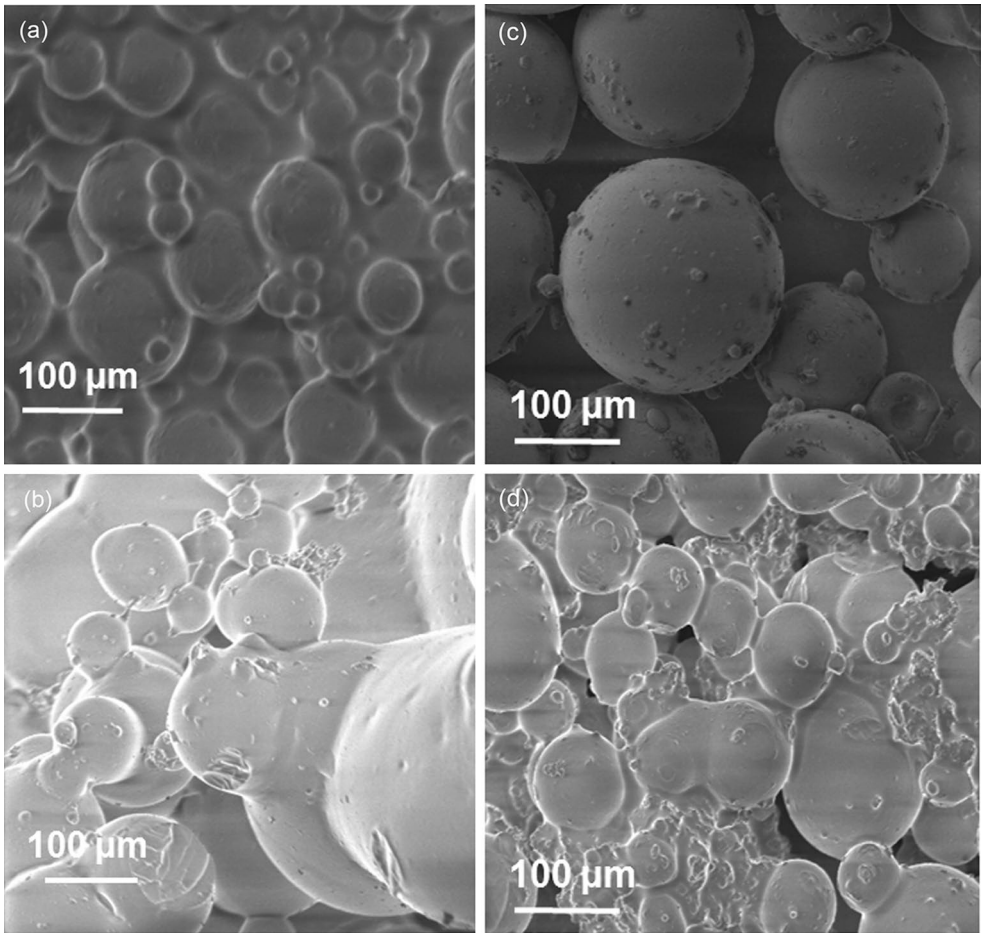


Fig. 4 SEM micrographs of microcapsules obtained at different temperatures: **a** 50 °C, **b** 60 °C, **c** 70 °C, and **d** 75 °C (stirring speed = 600 rpm)

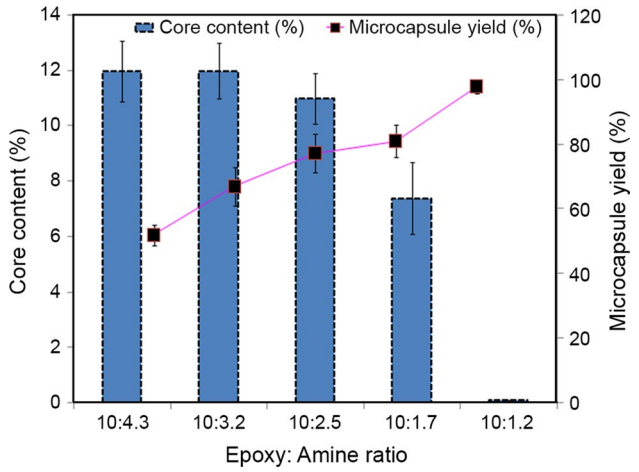


Fig. 5 Effect of resin/hardener ratio on core content and microcapsule yield

Effect of epoxy/amine ratio on yield and core content

The effect of varying the epoxy/amine ratio on the microcapsule yield and the amine core content is presented in Fig. 5. The core content is indicative of the amount of liquid amine actually available in the microcapsules for healing purposes, which was quantified as the ratio of mass loss during water reflux to original mass of sample taken.

As expected, microcapsule yield increased with increasing epoxy concentration until the stoichiometric ratio of 10/1.2 was reached. However, this is associated with a concomitant decrease in the core content. It is to be noted that microcapsules could not be obtained as the epoxy/amine ratio was reduced to less than 10/4.3. Although, attaining larger core content is a vital target of any encapsulation process, it was not possible to increase the core content further, probably due to the tacky nature of the microcapsules, which led to their agglomeration on increasing the amine content in the feed [31].

It is to be noted that in view of the miscibility of epoxy and amine (closeness of solubility parameter), the formation of "monolithic" epoxy microcapsules is expected [32]. SEM micrograph of a broken microcapsule is presented in the supplementary information file, which clearly reveals the presence of microdroplets of amine dispersed within the epoxy shell.

Effect of stirring speed on particle size distribution and morphology

An optical image of a representative batch of the free flowing microcapsules is presented in Fig. S4. Among the experimental parameters: stirring speed particularly, defines the equilibrium between shear forces and interfacial tension of the discrete phase-separated droplets [33]. The SEM micrographs of the microcapsules prepared by varying the stirring speed are presented in Fig. 6 and the associated particle size distributions are

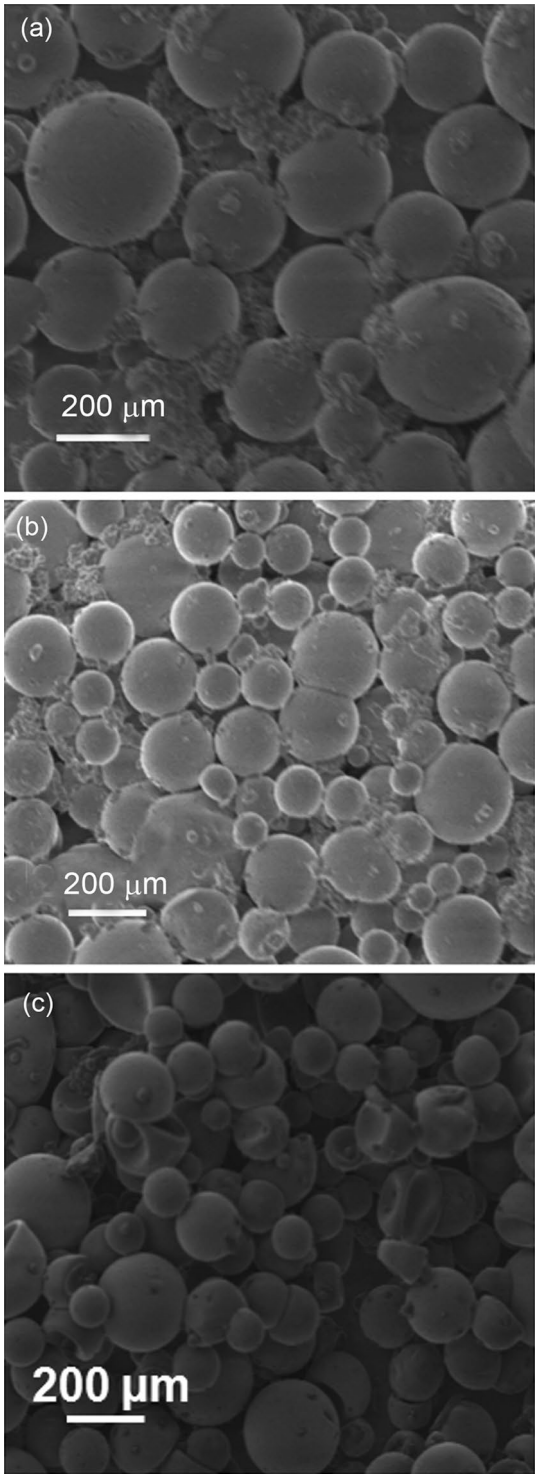


Fig. 6 Effect of stirring speed on the surface morphology of microcapsules: **a** 400, **b** 500, and **c** 600 rpm

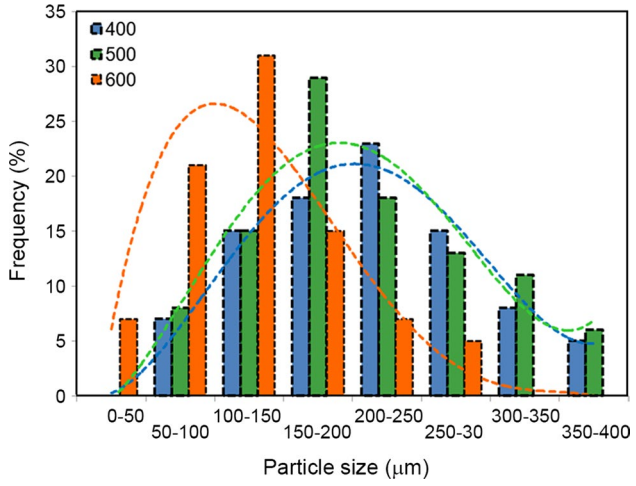


Fig. 7 Effect of stirring speed on the average particle size distribution of microcapsules

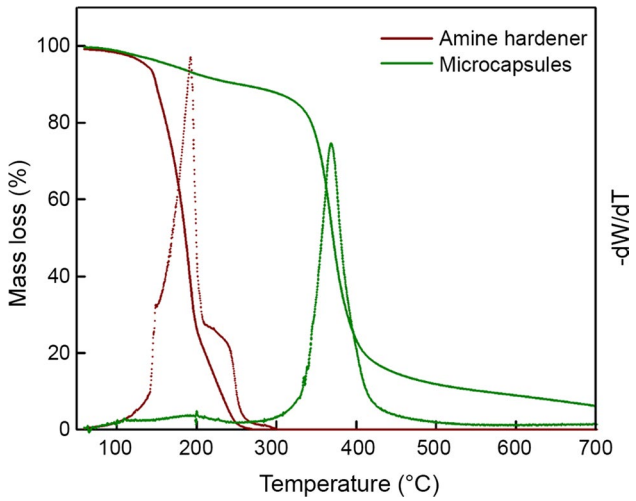


Fig. 8 TG-DTG thermograms of **a** TETA hardener and **b** microcapsules encapsulating TETA (*Dotted lines* represent the DTG thermograms)

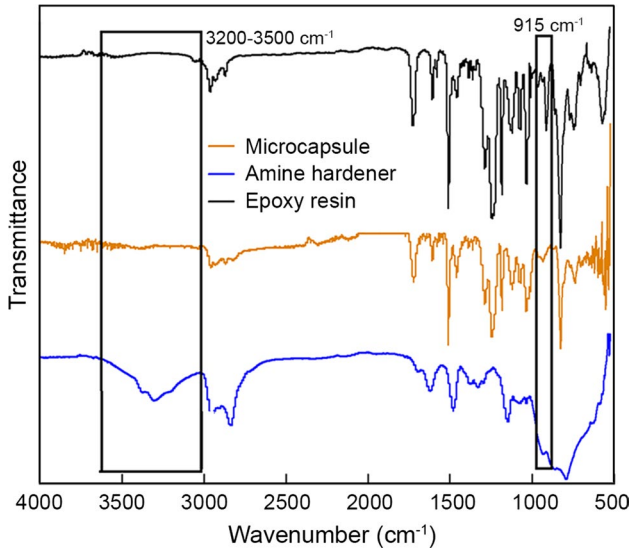


Fig. 9 FTIR spectra of **a** liquid amine and **b** microcapsules prepared using epoxy/hardener ratio = 10/3.2

presented in Fig. 7. As expected, increasing the stirring speed led to a shift in the average particle size towards lower dimensions, which could be attributed to the extensive shearing of the large oily droplets into smaller microspheres. It was observed that free flowing microcapsules could not be obtained upon lowering the stirring speeds to <400 rpm. Increasing the stirring speed to >600 rpm led to the formation of microcapsules of significantly lower dimensions which were of any practical significance. Under optimal reactions conditions, i.e., stirring speed of 400 rpm, reaction temperature of 70 °C and resin/hardener ratio of 10/3.2, spherical microcapsules could be obtained.

The TG-DTG thermograms of the liquid amine hardener before and after encapsulation are presented in Fig. 8. As can be seen, liquid amine volatilizes and undergoes oxidative degradation at $T \sim 210$ °C. As expected, the encapsulation of the amine in cross-linked epoxy shell shifts the degradation temperature towards higher temperature ($T_{onset} = 400$ °C).

The FTIR spectra of liquid epoxy resin and microcapsules are presented in Fig. 9. It can be seen that the intensity of the absorption band due to –C–O–C (oxirane ring 915 cm^{-1}) [34] reduces as a result of reaction with amine. It was not possible to determine epoxy content in the microcapsules prepared using 10/3.2 epoxy/hardener concentration.

Stability of the microcapsules

The stability of the microcapsules was established by placing them under isothermal conditions (30 °C) over extended time periods (3 months) and estimating the mass loss as a function of time. Negligible mass loss was observed over the entire period which clearly highlighted the potential of these microcapsules for practical applications.

Reactivity of the encapsulated amine

The curing studies of epoxy resin with encapsulated amine were also performed using non-isothermal calorimetry. The slurry formed after rupture of the microcapsules with epoxy was placed in a DSC pan and subjected to a controlled temperature program and the resulting curing profile is presented in Fig. 10. The profile of neat

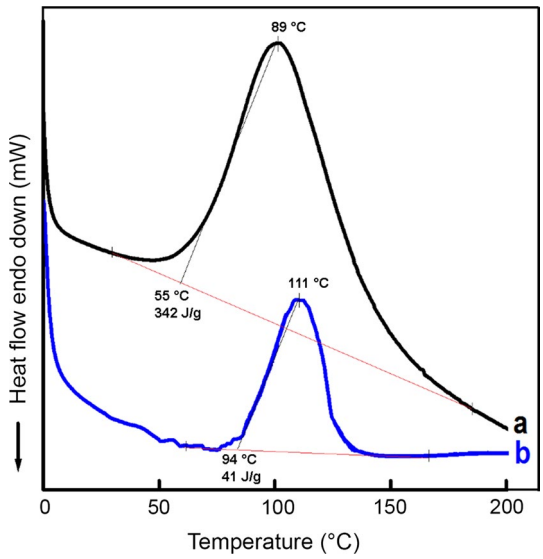


Fig. 10 Curing profile of epoxy with **a** amine and **b** crushed microcapsules

epoxy-amine is also presented for reference. It is apparent from the figure that the ring opening reaction shifts towards higher temperature in the case of microcapsule-epoxy slurry. In addition, the degree of heat evolved was ~12–13% of the heat associated with the curing of the neat epoxy: an observation in accordance with the core content determined gravimetrically.

To demonstrate the potential of the amine-filled microcapsules for healing applications, the microcapsules were crushed with a mortar and pestle and were added to neat epoxy resin CY230 resin at 1:1 mass ratio. The resulting slurry was placed between two glass slides and after resting for 72 h at room temperature; the slides could not be manually separated by shear, indicating the completion of curing reaction between the ruptured amine and the epoxy resin. In contrast, control slides with only epoxy resin or only amine microcapsules were easily separated and no film formation was evident.

Conclusion

Cross-linked epoxy microcapsules containing reactive amine hardener in their cores were prepared utilizing the principles of interfacial engineering. Reaction parameters, especially epoxy resin/hardener ratio, stirring rate, and temperature were found to have pronounced effect on the microcapsules morphology, particle size distribution and core content. Under optimal conditions, i.e., stirring speed of 600 rpm, reaction temperature of 70 °C and epoxy/hardener ratio 10/3.2, spherical microcapsules were obtained in excellent yields (100%) with a core content of 12%.

Acknowledgements The authors gratefully acknowledges DRDO for funding this research work. Thanks are also to Dr Parshuram Shukla for meaningful discussions throughout the research work (Grant No. ST/14-15/CFE-1311).

References

1. Blaiszik BJ, Kramer SLB, Olugebefola SC, Moore JS, Sottos NR, White SR (2010) Self-healing polymers and composites. Annu Rev Mater Res 40:179–211
2. Hillewaere XKD, Du Prez FE (2015) Fifteen chemistries for autonomous external self-healing polymers and composites. Prog Polym Sci 49–50:121–153
3. Zhu DY, Rong MZ, Zhang MQ (2015) Self-healing polymeric materials based on microencapsulated healing agents: from design to preparation. Prog Polym Sci 49–50:175–220
4. Toohey KS, Sottos NR, Lewis JA, Moore JS, White SR (2007) Self-healing materials with microvascular networks. Nat Mater 6:581–585
5. White SR, Sottos NR, Geubelle PH, Moore JS, Kessler MR, Sriram SR, Brown EN, Viswanathan S (2001) Autonomic healing of polymer composites. Nature 409:794–797
6. Jackson AC, Bartelt JA, Marczewski K, Sottos NR, Braun PV (2011) Silica-protected micron and sub-micron capsules and particles for self-healing at the microscale. Macromol Rapid Commun 32:82–87
7. Yang G, Mauldin TC, Lee JK (2015) Cure kinetics and physical properties of poly(dicyclopentadiene/5-ethylidene-2-norbornene) initiated by different Grubbs' catalysts. RSC Adv 5:59120–59130
8. Guadagno L, Raimondo M, Naddeo C, Longo P, Mariconda A (2014) Self-healing materials for structural applications. Polym Eng Sci 54:777–784
9. Kamphaus JM, Rule JD, Moore JS, Sottos NR, White SR (2008) A new self-healing epoxy with tungsten (VI) chloride catalyst. J R Soc Interface 5:95–103
10. Cho SH, Andersson HM, White SR, Sottos NR, Braun PV (2006) Polydimethylsiloxane-based self-healing materials. Adv Mater 18:997–1000
11. Li C, Tan J, Gu J, Qiao L, Zhang B, Zhang Q (2016) Rapid and efficient synthesis of isocyanate microcapsules via thiol-ene photopolymerization in Pickering emulsion and its application in self-healing coating. Compos Sci Technol 123:250–258
12. Keller MW, White SR, Sottos NR (2007) A self-healing poly(dimethyl siloxane) elastomer. Adv Funct Mater 17:2399–2404
13. Yuan YC, Rong MZ, Zhang MQ, Chen J, Yang GC, Li XM (2008) Self-healing polymeric materials using epoxy/mercaptan as the healant. Macromolecules 41:5197–5202
14. Xiao DS, Yuan YC, Rong MZ, Zhang MQ (2009) Self-healing epoxy based on cationic chain polymerization. Polymer 50:2967–2975
15. Tripathi M, Dwivedi R, Kumar D, Roy PK (2016) Application of microencapsulated unsaturated polyester toward temperature-triggered healing in epoxy composites. J Intell Mater Syst Struct 27:1650–1657
16. Yin T, Rong MZ, Zhang MQ, Yang GC (2007) Self-healing epoxy composites—preparation and effect of the healant

consisting of microencapsulated epoxy and latent curing agent. Compos Sci Technol 67:201–212
17. Yuan L, Liang G, Xie JQ, Li L, Guo J (2006) Preparation and characterization of poly(urea-formaldehyde) microcapsules filled with epoxy resins. Polymer 47:5338–5349
18. Cosco S, Ambrogi V, Musto P, Carfagna C (2006) Urea–formaldehyde microcapsules containing an epoxy resin: influence of reaction parameters on the encapsulation yield. Macromol Symp 234:184–192
19. McIlroy DA, Blaiszik BJ, Caruso MM, White SR, Moore JS, Sottos NR (2010) Microencapsulation of a reactive liquid-phase amine for self-healing epoxy composites. Macromolecules 43:1855–1859
20. Kobašlija M, McQuade DT (2006) Polyurea microcapsules from oil-in-oil emulsions via interfacial polymerization. Macromolecules 39:6371–6375
21. Poe SL, Kobašlija M, McQuade DT (2006) Microcapsule enabled multicatalyst system. J Am Chem Soc 128:15586–15587
22. Jin H, Mangun CL, Stradley DS, Moore JS, Sottos NR, White SR (2012) Self-healing thermoset using encapsulated epoxy–amine healing chemistry. Polymer 53:581–587
23. Zhang H, Wang P, Yang J (2014) Self-healing epoxy via epoxy–amine chemistry in dual hollow glass bubbles. Compos Sci Technol 94:23–29
24. Sharma P, Lochab B, Kumar D, Roy PK (2015) Interfacial encapsulation of bio-based benzoxazines in epoxy shells for temperature triggered healing. J Appl Polym Sci 132:42832. doi:10.1002/app.42832
25. Ullas AV, Rahamattulah Kumar D, Roy PK (2016) Epoxy filled microcapsules by interfacial engineering. Polym Plast Technol Eng 55:937–942
26. Brown EN, Kessler MR, Sottos NR, White SR (2003) In situ poly(urea-formaldehyde) microencapsulation of dicyclopentadiene. J Microencapsul 20:719–730
27. Ramanathan LS, Baskaran D, Shukla PG, Sivaram S (2002) Preparation of polyurethane microspheres via dispersion polycondensation using poly(1,4-isoprene)-block-poly(ethylene oxide) as steric stabilizer. Macromol Chem Phys 203:998–1002
28. Hoy KL (1970) New values of the solubility parameters from vapor pressure data. J Paint Technol 42:76–118
29. Budrugeac P, Segal E (2007) Applicability of the Kissinger equation in thermal analysis. J Therm Anal Calorim 88:703–707
30. Tripathi M, Kumar D, Rajagopal C, Roy PK (2015) Curing kinetics of self healing epoxy thermosets. J Therm Anal Calorim 119:547–555
31. Fiedler B, Gojny FH, Wichmann MHG, Nolte MCM, Schulte K (2006) Fundamental aspects of nano-reinforced composites. Compos Sci Technol 66:3115–3125
32. Tripathi M, Kumar D, Roy PK (2017) Healant release from microcapsules with varied internal microstructure. Compos Commun 3:38–41
33. Rallison JM (1984) The deformation of small viscous drops and bubbles in shear flows. Annu Rev Fluid Mech 16:45–66
34. Ramírez C, Rico M, Torres A, Barral L, López J, Montero B (2008) Epoxy/POSS organic–inorganic hybrids: ATR-FTIR and DSC studies. Eur Polym J 44:3035–3045

2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing

Optimized MapFile based Storage of Small files in Hadoop

Divyashikha Sethia, Shalini Sheoran
Computer Engineering Department
Delhi Technological University
Delhi, India
Email: sethiadivya@gmail.com,
shalini100sheoran@gmail.com

Huzur Saran
Computer Science & Engineering Department
IIT Delhi
Delhi, India
Email: saran@cse.iitd.ernet.in

***Abstract*—Hadoop is an open source software based on MapReduce framework. The Hadoop Distributed File System (HDFS) performs well while storing and managing data sets of very large size. However, the performance of HDFS suffers while handling a large number of small files since they put a lot of burden on the NameNode of HDFS both in terms of memory and access time. To overcome these defects, we merge small files into a large file and store the merged file on HDFS. Generally, when small files are merged, variation in the size distribution of files is not taken into consideration. We propose a new algorithm OMSS (Optimized MapFile based Storage of Small files) which merges the small files into a large file based on the Worst fit strategy. The strategy helps in reducing internal fragmentation in data blocks, which in turn leads to fewer data blocks consumed for the same number of small files. Less number of data blocks mean fewer memory overheads at major nodes of Hadoop cluster and hence increased efficiency of data processing. Our experimental results indicate that the time to process data on HDFS containing unprocessed small files reduces significantly to 590s when MapFile is used and it reduces further to 440s when OMSS is used. OMSS as compared to MapFile merging algorithm has a reduction of 34.7% in memory requirements.**

***Index Terms*—Hadoop; HDFS; MapFile; OMSS; NameNode**

I. Introduction

In the present world of cloud computing, a huge amount of data is being generated and this data needs to be stored, processed and analyzed. HDFS [1] serves as primary storage system of Hadoop and has a master-slave architecture. HDFS is a common representative for Cloud service filesystem running on clusters and has been widely used to support Cloud applications. Having on-demand availability to computation, high scalability and large storage capacity, cloud computing [2] is the ideal solution for big data analytics and processing. Deploying Hadoop on Cloud eliminates the operational challenges of running Hadoop. HDFS is suitable for the analysis of large datasets of applications such as machine learning, data mining, etc. But, HDFS ignores the problem associated with access and storage of small files. In applications like meteorology, energy, e-Library, messaging apps, e-Business, and e-Learning [3], high-performance computing files and healthcare records, data generated is in huge numbers, but, of very small sizes, typically ranging from 10KB to 10 MB. Datasets of digitized clinical data and genomic data records that helps researchers in performing data analytics on healthcare data are all stored in the form of small files.

Presently there are no available standard and generic file systems which are specifically designed for storing, processing and analyzing small files. Facebook has also faced a similar problem; it has about 15 billion photos [4] of its users growing at a rate of 220 million/week.For storing photos, Facebook has developed Haystack a distributed system to serve their specific needs. All currently available local file system, distributed file system, and the object-based storage systems such as Ceph, GFS, Google File, EXT4 are developed and designed for large files. Their performance degrades significantly in the case of processing and storing small files.Moreover, among the distributed system at the helm presently, Hadoop is prime big data analytics platform in the present data world.It is the leading platform regarding performance, reliability, scalability.

Thus, there are various reasons contributing to Hadoop small file problem [5]. Firstly, the NameNode memory footprint is a big issue. Every block, file, and directory in Hadoop is represented as an object in the memory on NameNode. As a rule, each small file requires 150 bytes of memory for storing its metadata. By reducing the number of small files on Hadoop cluster, we can reduce the NameNode memory consumption. Also, access time and network IO time are reduced as the number of requests to NameNode reduces with less number of small files. Secondly, the presence of a large number of small files will deteriorate the performance of MapReduce processing as a huge number of small files results in a large volume of random disk IO. For MapReduce performance, often disk IO is one of the prime limiting factors. Reading the data from, one large merged sequential file will always take less time as compared to reading the same size of data through multiple random reads. If we are able to store data in less number of blocks, the performance impact due to disk IO will be significantly mitigated. Practically, the generation of small files has become increasingly common. To solve the small file problem, we mainly merge small files and store the large file after merging. But, presently available solutions have not considered the size distribution of files and internal fragmentation caused by merging of files into consideration. In this paper, we firstly analyze, compare and contrast various

978-1-5090-6611-7/17 $31.00 © 2017 IEEE
DOI 10.1109/CCGRID.2017.83

IEEE computer society

solutions available for solving Hadoops small files problem, including Hadoops own solutions and other solutions which target various aspects of small files problem. In this paper, we propose a novel algorithm named as, OMSS (Optimized MapFile based Storage of Small files) to improve storage efficiency and processing time for small files in Hadoop. The OMSS technique reduces memory wastage done due to internal fragmentation by using Worst fit technique for merging several files into a MapFile [10]. We have implemented and compared the results on a Hadoop setup on cloud and find that our proposed algorithm performs better than basic MapFile technique.

Our solution can be applied to various big data applications such as in healthcare industry, where huge number of Electronic health records(EHR) are stored in healthcare information systems. Also, there are millions of users of various messaging apps, wherein text messages constitute a major chunk of data and this data is generally backed up every few minutes. It can be used in big data analytics and gaining insights in huge E-Commerce data.

The rest of this paper is organized as follows. Section II discusses background on Hadoop and Problems associated with small files; Section III describes related work; Section IV describes our proposed solution, OMSS; Section V contains experimental results and Section VI concludes the paper.

II. Background

In this section, we discuss architecture of Hadoop Distributed Filesystem (HDFS) and then the problems associated with processing, storing and accessing small files.

A. HDFS

HDFS [1] is the primary storage system of Hadoop and has a master/slave architecture. It consists of NameNode/s and DataNodes as architectural components. HDFS is one of the file systems of Hadoop written in Java which divides a large file into small blocks and distributes them to various nodes called DataNodes. Redundant copies of each block are maintained in HDFS to avoid loss of data in case of failures. NameNode maintains a database containing a mapping from logical files to physical blocks in the DataNodes. Block replication is also managed by the NameNode. The NameNode handles the management of the filesystem namespace, metadata, and requests from clients to access data. The DataNodes provide block storage and serve IO requests from clients. They also create, delete, and replicate data blocks upon getting command from the NameNode.

B. Problems with small Files

The HDFS is originally designed for storing and processing large files. We can consider a small file as any file that is considerably smaller than the Hadoop block, i.e., 64 MB. However, the small file problem is not just caused by the small size of files, if a large number of files in Hadoop cluster are marginally greater than an increment of the block size, the user will encounter the same challenges as small files. In present day scenario with the help of cloud computing, everybody wants data to be available in real time. Due to real time syncing of data, Hadoop ingestion is made to run frequently, which in turn leads to the creation of small files in HDFS.

HDFS has problems in processing and handling small files as mentioned below:

1) NameNodes High memory consumption: Metadata in stored in main memory of NameNode. In general, the metadata of a file consumes about 150 bytes of main memory. With say, 24 million files in HDFS, NameNode will require 3.5 GB of main memory for storing metadata. Therefore, an application consisting of huge number of small files will eventually consume all memory spaces of NameNode. Hadoop framework represents, every block, file, directory as an object in the memory.So, If we are able to reduce the number of small files on Hadoop cluster, we can also reduce network impact and startup time along with the NameNode memory footprint.

2) The MapReduce performance problem: Having a large number of small files degrades the performance of MapReduce [6] processing. The reason for performance degradation is that a huge number of small files means a significant amount of random disk IO, which is often one of the prime limiting factors in MapReduce performance. Reading the same data in a single sequential read will generally take less time than reading via multiple random reads. We can reduce this performance penalty due to random disk IO if we can store our data on few large blocks.

3) High storing time: Storage time of small files in HDFS is quite high. For example [7], to store 550,000 small files into HDFS system, it requires around 7.7 hours. The files in this case vary in size from 1KB to 10KB. On the other hand, if we attempt to store these files using a local file system, it takes considerably less time. Thus showing unfeasibility of storing time.

4) NameNode performance degradation: HDFS takes a significant amount of time for managing metadata as it requires nodes to cooperate. When HDFS client receives a request for accessing a file it needs to retrieve metadata of the file from Namenode for each such request for file access. For small files, the major overheads are metadata management and disk seeks. While data transfer takes very short time. HDFS client has to contact NameNode frequently and as a result performance of NameNode is impacted, when there are huge number of small files in system.

III. Related Work

In recent times, various organizations using Hadoop as well as research community have given a lot of attention towards small file problem.

Some of the general solutions for solving small file problem are:

HAR : HAdoop Archive (HAR) [8] archives small files into files using Hadoop archive method. Hadoop archive method uses MapReduce to merge small files into fewer large HDFS files. It contains metadata files and data files.The metadata file contains a index and master index. Index file has the

name and location of files that are part of an archive. HAR is effective in reducing memory consumption of NameNode. The data file consists of files which are archived together. But there are many drawbacks of HAR, each access to HAR file requires 2 reads from the index file and one read from the data file. And also, new files cannot be appended to HAR, they are immutable once created. We need to re-create HAR for inserting new files.

SequenceFile : SequenceFile [9] is a data structure which stores the data in the form of binary key-value pairs. Also, it acts as a container for small files. Here key is file name and value is file contents. Compression and decompression are supported both at record level and block level. Sequence File also has several disadvantages. For a particular key, it does not have mechanism for update and delete operation; it only supports append method; and secondly, this approach has low access efficiency as it takes quite a long time to make a Sequence file. If a user needs to look up for a particular key, the complete sequence file is required to be read.

MapFile: A MapFile [10] is a sorted Sequence File. It comprises of two files, one index file and another one is data file. It maintains index file to store key location information to allow lookup of data by key. The key-value pairs are sorted by key and stored as records in the data file. MapFile facilitates to look up for key without needing to read full file. Drawbacks of MapFile are similar to Sequence File. They cannot provide flexible APIs for applications as only append method is supported for a particular key. Moreover, while storing files correlations between files are not considered in all three methods namely, HAR, SequenceFile, and MapFile.

Various authors have also proposed application specific and several domain specific solutions for small file problems. L.Xuhui Liu, et al. [7] proposed an approach where the main idea was to combine small files into large ones to reduce the file number and build an index for each file. Additionally, they provided some features such as grouping neighboring files and reserving several latest version of data considering the characteristics of WebGIS access patterns. Their technique was similar to MapFile. They did not focus on size distribution and processing time of small files.

G. Mackey et al.[11] proposed a scheme based on HAR file layout to solve the small files problem. Their primary focus was on the scheduling of jobs on Hadoop cluster using quota policy. The HAR file layout itself has several disadvantages as mentioned above in this paper. They did not introduce any good solutions for reducing access time, processing time and memory consumption at major nodes of Hadoop.

B. Dong et al. [12] presented a scheme in that, firstly, it merges all correlated small files of a PPT courseware into a larger file to reduce the memory footprint on NameNode. Secondly, to reduce the access time of small files, it uses a two-level prefetching mechanism.

The authors in their subsequent research [13], introduced an approach which focussed on improving the storage capabilities and reducing access time of small files on HDFS. File merging used by them is similar to other techniques. They classified files as structurally-related files, logically-related files, and independent files. And a prefetching scheme was applied for structurally-related small files, while file grouping and prefetching scheme were used for managing logically-related small files. Both the approaches proposed by authors were very similar to MapFile technique. In their work, they considered correlation among small files and tried to identify the relationship among various small files. But, here also size distribution of files was not considered while merging small files. Their prime focus was on the needs of their real education system.

To further improve the storing capability and accessing time of the small files on HDFS, Chandrasekar S et al. [14] proposed a solution based on the works of Dong et al., namely Extended Hadoop Distributed File System (EHDFS). Here, in this scheme, correlated files are identified by the client and combined into a single large file to minimize total count of files. An indexing mechanism was created to access the individual files from the corresponding combined file. Further, index prefetching was also provided to improve IO performance and minimize the load on NameNode.

In this paper, we propose an Optimized MapFile based algorithm for Storage of Small files, considering the size distribution of files. The concept of considering size distribution of files for merging small files was proposed by Hue He et al. The solution presented by Hue He et al. [15] was a Tetris merge algorithm which was based on the balance of data block. The merging strategy was similar to filling of gaps in Tetris. In the proposed scheme, based on the volume of small files, the small files are evenly distributed into bigger files. Taking the size distribution of small files into consideration, we have proposed an OMSS algorithm. In our proposed algorithm, we have modified the merging process of MapFile in order improve storage and processing efficiency. We have used worst fit strategy to merge files into MapFile. Moreover, using OMSS, we have been able to reduce processing time and memory burden on the primary nodes of HDFS. Hence using OMSS we can resolve the small files problem significantly and benefit with various applications such as healthcare and education which host large, small sized data by providing fewer memory requirements and faster processing time.

IV. Proposed Work

In this paper, we have proposed a modified MapFile technique which attempts to reduce memory wastage while merging small files into large MapFiles. To overcome problems in handling small files, we merge small files into a larger file and store the merged file on HDFS. While merging small files, variation in the size distribution of files is not taken into consideration. We propose a new algorithm OMSS (Optimized MapFile based Storage of Small files) which takes into account size distribution of files present in a file set. It merges the small files into larger files based on Worst fit strategy that helps in reducing internal fragmentation in data blocks, which in turn leads to fewer data blocks consumed for the same number of small files. Less number of data blocks means fewer memory

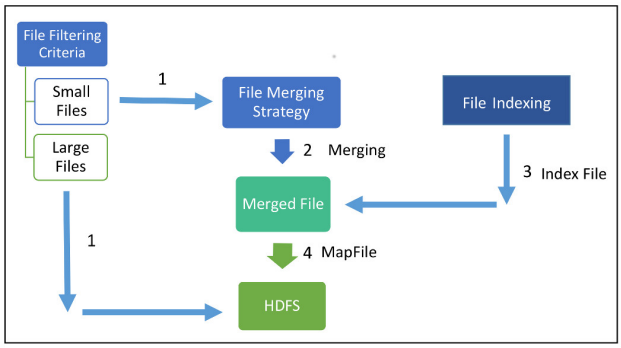


Fig. 1. Schematic Diagram of Proposed solution

overheads at major nodes of Hadoop cluster and increased efficiency of data processing.

In our proposed algorithm we have reduced internal fragmentation in MapFile using Worst-Fit strategy. In the worst-fit strategy, process is placed in the biggest block of unallocated memory available by memory manager. As we have mentioned earlier small file problem is not just due to very small files, but it is also created by files that are smaller than block size but due to less space available in the data block, they spill over to next data block of DataNode and create the small file problem. As mainly, all the merging algorithms are based on the cut-off size of the merged file, while merging many small files, if the size of the large file being created reaches the cut-off point, the files are merged into a bigger file. In this process if a file crosses cut off size, it is removed from that large file, and internal fragmentation occurs in the merged file. As shown in fig. 2, while merging 3 files i.e. 1,2,3 total sizes of these 3 files crosses the cut-off size pre decided by a merging algorithm, then file 3 is removed from the file being merged. And thus leaving behind memory unutilized. Hence it is imperative to take into account size distribution of files and merging files accordingly; By this approach we can optimize memory requirements of major nodes of Hadoop cluster.

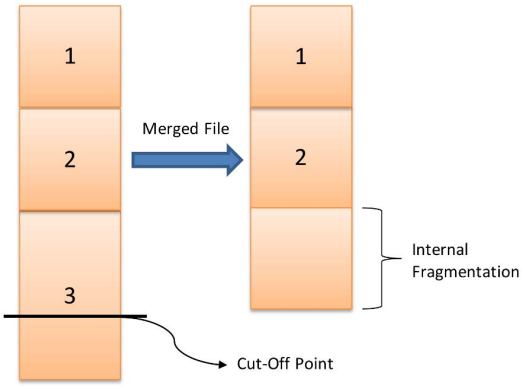


Fig. 2. Problem of Internal Fragmentation

The File Set contains the set of files to be processed into Hadoop Cluster.As shown in fig. 1, we first segregate large files present in the File Set from small files. Then the large files are directly put into Hadoop Cluster for processing. Using the worst fit strategy of file merging, as proposed in our algorithm, we sort the small files in the File Set into various merging queues. Small files are put into the queue with the maximum merge limit. Files are placed into the merging queue until the queue size become equal to merge criteria. After the files are sorted into various queues, the files are merged and converted into MapFiles, one per queue. MapFile is a file containing sorted key-value pairs. Here key is file name and value is file contents. MapFile also maintains an index file to facilitate faster access of small files. These MapFiles are then put into Hadoop Cluster for processing.

Our proposed File merging strategy for MapFiles consists of two parts (1) File filtering based on the size of small files, (2) File merging using worst-fit strategy.

1) File Filtering Criteria: For optimizing performance of HDFS for handling small files, it is crucial to decide the cut off point between small files and large files. Many applications have a huge number of files, which are considerably small in size. For example, in the field of climatology [16], some applications contains 450,000 files with an average size of 61MB. In Sloan Digital Sky Survey, there were around 20 million image files [17] with an average size even less than 1MB. Although Dong et al. discussed the cut-off point between large and small files, they just presented the experimental analytical results.

Algorithm 1 File Merging Algorithm

1: Initialize Fileset(Fs), MergeCriteria(Ms), MergeQueue(Qm)
2: **for** each file F_i in Fs **do**
3: **if** (F_i belongs to Fs and F_i is not a small file) **then**
4: Goto Step 3
5: **else**
6: Goto Step 7
7: Merging of the files using worst fit strategy where merge limit is the maximum limit. Select the queue with the max merge limit i.e. max free size
8: Insert the file into selected queue and calculate the queue size(Qs) and merge limit as:
Qs = Qs + file size(F_i)
merge limit = Ms - Qs
9: **if** (Qs==Ms) **then**
10: Make a sequence file of the files in the queue
11: **else**
12: Goto Step 2
13: Put all the sequence files generated and large files present in fileset into HDFS for processing.

2) File Merging: We can understand file merging process as shown in fig. 3, there are five files namely 1,2,3,4,5 in the file list. Out of which File 3 represents a large file, while all

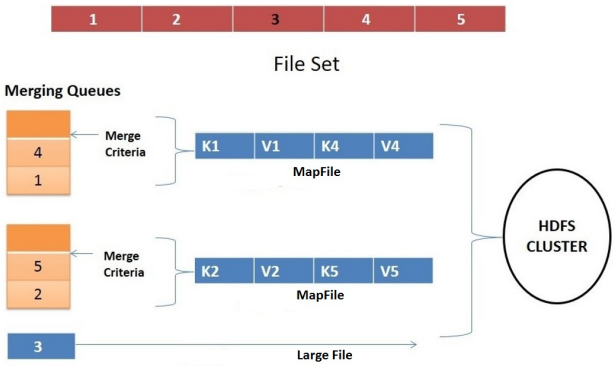


Fig. 3. Schematic Diagram of File Merging Strategy

other are small files. Here, sizes of file 1,2,3,4,5 are taken 6,7,200,4,2 units respectively. And let's consider threshold limit to be 10 units. Our file merging algorithm puts the file into various merging queues based on worst fit strategy. File 1 and 4 are placed into the 1st queue, while 2 and 5th into the second one. The files in the queue are then converted into MapFile (key, value pair). The MapFiles generated and the large file are then put into HDFS cluster for processing.

As described in algorithm 1, We first initialize the set of all input files(Fs), the criteria for merging the files (Ms) and the merging queue (Qm). Next, we iterate for every input file of our set, if the file is a large file, we put it into the Hadoop Distributed File system for its indexing, processing, and storage in the cluster. Else, if the file is indeed a small file, then we merge it into the Merge Queue, using the worst fit strategy, that is, we put it into the queue having the maximum possible empty space. After merging the file into the merge queue, its queue size is incremented by the file size, and the merging limit is decremented by the queue size. Whenever, for a queue, the queue size becomes equal to the merge size limit, the queue is converted into a sequence file. And the sequence files are put in the Hadoop DFS for processing. This algorithm makes use of the worst fit strategy, to reduce internal fragmentation in storing of files. As the hole (empty space) which is left after inserting a file into a queue is the maximum, hence the chances of accommodating any new file into the left over space are maximum. Thereby wastage of space is minimized.

V. Experimental Evaluation

A. Experimental Environment

The experimental setup is established on a cloud named as Baadal [18]. Baadal: is a IITD cloud computing environment for high-performance computing. It is an indigenously developed cloud orchestration and virtualization management software. Baadal has the capability to work with multiple virtualization technologies like KVM, Xen and VMware. Main capabilities of Baadal are:

- Dynamic resource scheduling and power management.
- An integrated work flow system for requesting and commissioning of virtual machines.
- Monitor resource utilization dynamically.
- Facilities for specifying resource requirement, power on-off, shut down, suspending and resuming of virtual machines.
- A minimal cost model for virtual machine resources.

TABLE I
CONFIGURATION OF ENVIORNMENTAL SETUP

Memory	2GB
Hard Disk	500GB
Operating System	Ubuntu 12.04
CPU	2.6GHz
Hadoop Version	2.6.0
Connectivity	10Gbps

Baadal is deployed in IIT-Delhi on 48 blade servers with around 500 cores and virtualized storage of 50 TB. It has currently more than 60 virtual machines provisioned for high-performance computing.

The setup on Baadal consists of four nodes, having one node as NameNode and three other nodes as DataNodes. Each node of the cluster has Ubuntu 12.04 installed on it and has Hadoop 2.6.0 version. The Java version of operating environment is 1.6. All four nodes are Intel CPU (2.6 GHz), 2GB memory and 500 GB disk, and act as DataNodes and NameNode.

For conducting experiments on Hadoop cluster deployed on Baadal, a total 3420 files of different sizes have been used. Such as 2000 files of less than 100KB, then 300 files of size less than 500KB, 1000 files of less than 1MB size,100 files of less than 64 MB and few files larger than 64MB.

B. Experimental Results

1) Comparison of memory consumption at Namenode: The experiment was carried out both using MapFile and OMSS techniques. For conducting experiments on the Hadoop cluster, the data set consisted of 3420 files of varying sizes stored on baadal, making total size of data to be 2.4GB. When all files were put directly into HDFS without any preprocessing, the HDFS considered each file as an object and allocated each file a data block at DataNode. Each data block consumed 150 bytes of memory at NameNode. So for 3420 files, each consuming 150 bytes of data, a total of 500.97 KB data was consumed at NameNode. When MapFile was used for merging small files only 10.54KB of data was consumed at NameNode. And when OMSS was applied then memory consumption was further reduced to 6.88KB.

TABLE II
MEMORY CONSUMPTION AT NAMENODE

Technique used	Memory consumption(KB)	Data blocks needed for merging
HDFS	500.97	3420
MapFile	10.54	72
OMSS	6.88	47

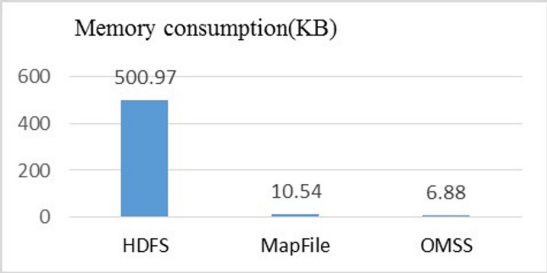


Fig. 4. Memory consumption at NameNode

2) Comparison of data processing time: After merging files, memory consumption at NameNode was reduced. Also, the processing time of MapFile by Hadoop's MapReduce framework was reduced considerably. Here processing time includes both map and reducce stages. For calculating processing time of small files we used word count problem. We separately processed the files put in HDFS using all three methods i.e. HDFS (without any preprocessing), Mapfile, OMSS. As data in Table III show that time to process data on HDFS containing unprocessed small files reduces significantly to 590s when MapFile was used and it reduces further to 440s when OMSS was used.

TABLE III
DATA PROCESSING TIME AT NAMENODE

Technique used	Time consumed (in s)
HDFS	4560
MapFile	590
OMSS	440

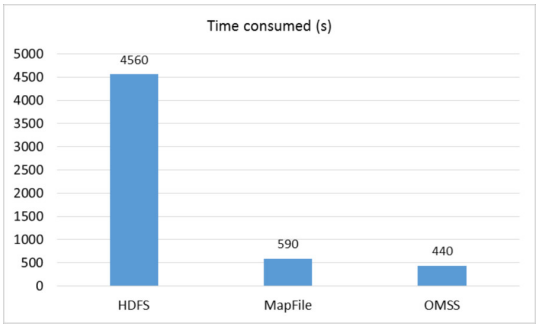


Fig. 5. Data processing time at NameNode

VI. Conclusion

With this paper, we have proposed a novel method which is an improvement over MapFile by modifying the file merging method. We have provided an optimized strategy for file merging and storage by considering drawbacks of Hadoop distributed file system for processing and storage of small files. The method is to organize files into optimized groups using the worst fit strategy of file merging by using the distribution of volume of files, and it has significantly reduced the metadata size of NameNode and memory consumption at HDFS. Our experimental results show that the proposed method has also improved the data processing efficiency of MapReduce. The time to process data on HDFS containing unprocessed small files reduces significantly to 590s when MapFile is used, and it reduces further to 440s when OMSS is used. As compared to MapFile merging algorithm, OMSS has a reduction of 34.7% in terms of memory requirements. Using OMSS we can resolve the small files problem significantly and benefit with various applications such as healthcare and education which hosts large, small sized data by providing less memory requirements and faster processing time.

As a part of future work, the problem can be considered as an online problem that can speculatively append files based on size to limit fragmentation, instead of current offline methods which require first all files to be available for future processing of files. And also the decision to append can (which files and where) be probabilistic based on observations about cluster utilization.

References

[1] Hdfs architecture guide. [Online]. Available: https://hadoop.apache.org/docs/r1.2.1/hdfs_design.html
[2] P. Mell and T. Grance, "A nist definition of cloud computing. national institute of standards and technology. nist sp 800-145," 2009.
[3] B. Dong, Q. Zheng, M. Qiao, J. Shu, and J. Yang, "Bluesky cloud framework: an e-learning framework embracing cloud computing," in *IEEE International Conference on Cloud Computing*. Springer, 2009, pp. 577–582.
[4] M. Marcon, B. Viswanath, M. Cha, and K. P. Gummadi, "Sharing social content from home: a measurement-driven feasibility study," in *Proceedings of the 21st international workshop on Network and operating systems support for digital audio and video*. ACM, 2011, pp. 45–50.
[5] T. White. The small files problem. [Online]. Available: http://blog.cloudera.com/blog/2009/02/the-small-files-problem/
[6] J. Dean and S. Ghemawat, "Mapreduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107–113, 2008.
[7] X. Liu, J. Han, Y. Zhong, C. Han, and X. He, "Implementing webgis on hadoop: A case study of improving small file i/o performance on hdfs," in *2009 IEEE International Conference on Cluster Computing and Workshops*. IEEE, 2009, pp. 1–8.
[8] Hadoop archive guide. [Online]. Available: https://hadoop.apache.org/docs/r1.2.1/hadoop_archives.html
[9] Sequence file. [Online]. Available: https://wiki.apache.org/hadoop/SequenceFile
[10] Mapfile. [Online]. Available: https://hadoop.apache.org/docs/r2.6.0/api/org/apache/hadoop/io/MapFile.html
[11] G. Mackey, S. Sehrish, and J. Wang, "Improving metadata management for small files in hdfs," in *2009 IEEE International Conference on Cluster Computing and Workshops*. IEEE, 2009, pp. 1–4.
[12] B. Dong, J. Qiu, Q. Zheng, X. Zhong, J. Li, and Y. Li, "A novel approach to improving the efficiency of storing and accessing small files on hadoop: a case study by powerpoint files," in *Services Computing (SCC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 65–72.
[13] B. Dong, Q. Zheng, F. Tian, K.-M. Chao, R. Ma, and R. Anane, "An optimized approach for storing and accessing small files on cloud storage," *Journal of Network and Computer Applications*, vol. 35, no. 6, pp. 1847–1862, 2012.
[14] S. Chandrasekar, R. Dakshinamurthy, P. Seshakumar, B. Prabavathy, and C. Babu, "A novel indexing scheme for efficient handling of small files in hadoop distributed file system," in *Computer Communication and Informatics (ICCCI), 2013 International Conference on*. IEEE, 2013, pp. 1–8.

[15] H. He, Z. Du, W. Zhang, and A. Chen, "Optimization strategy of hadoop small file storage for big data in healthcare," *The Journal of Supercomputing*, pp. 1–12, 2015.
[16] A. Chervenak, J. M. Schopf, L. Pearlman, M.-H. Su, S. Bharathi, L. Cinquini, M. D'Arcy, N. Miller, and D. Bernholdt, "Monitoring the earth system grid with mds4," in *2006 Second IEEE International Conference on e-Science and Grid Computing (e-Science'06)*. IEEE, 2006, pp. 69–69.
[17] E. H. Neilsen Jr, "The sloan digital sky survey data archive server," *Computing in Science & Engineering*, vol. 10, no. 1, pp. 13–17, 2008.
[18] Baadal: the iitd computing cloud. [Online]. Available: http://www.iitd.ac.in/content/baadal-iitd-computing-cloud

Performance Analyses and Optimisation of Energy Recovery Systems for Marine Applications

O P Sharma[1*], S C Kaushik[2], K Manjunath[3]

[1,2]Centre for Energy Studies, Indian Institute of Technology, Delhi, India-110016

[3]Department of Mechanical Engineering, Delhi Technological University, Delhi, India-110042

*Corresponding author: Tel.: +919654671567, E-mail address: opsharma.iitd@gmail.com; kaushik@ces.iitd.ac.in

Synopsis

Naval ships are increasingly preferring gas turbines, over conventional engines such as diesel engines and steam turbines, as primemovers for propulsion as well as power generation owing to many advantages such as greater power density, compactness, flexibility of modular designs, lower weight and low noise signatures. However, gas turbines have relatively lower thermal efficiency. Therefore, the innovative measures for improvement of overall energy efficiency of such plants have assumed greater significance to the navies worldwide. Waste heat recovery of gas turbine exhaust to produce additional power is an attractive option to enhance thermal efficiency and reduce environmental impact. The supercritical CO_2 Brayton power cycle is a potential candidate for such waste heat recovery for naval applications due to its least volume and low weight. Such cycles have been extensively researched for heat recovery for advanced nuclear reactors. Therefore, in this paper, the authors have attempted to integrate a simple cycle marine gas turbine with supercritical CO_2 Brayton power cycle. A comparative energy and exergy based performance analyses and optimisation of a regenerative-Brayton cycle (RBC) and a regenerative-recompression-Brayton cycle (RRCBC), driven by the exhaust gas is presented. Results show that shipboard propulsion plant thermal efficiency enhances from 35% to 44.3% (RBC) and 45.1% (RRCBC) and its net work output increases by 26% (RBC) and 25% (RRCBC) at 100% load. The study will contribute in designing efficient yet compact waste heat recovery systems for shipboard applications.

Key words: waste heat recovery system; supercritical carbon dioxide; exergy; performance optimization; marine gas turbine; exhaust gas.

1. Introduction

Energy efficiency plays a critical role in limiting world energy demand growth to one-third by 2040, while global economy grows by 150% **[1]**. Amongst available technologies, waste heat recovery is the most effective yet self-sustainable and greener option to improve energy efficiency of any conventional thermal system. World over naval ships are increasingly preferring high speed diesel engine or gas-turbine based power plants due to features such as greater power density, compactness, flexibility of modular design, lower weight implication and lower noise signatures. However, these engines, despite recent improvements in their energy efficiency up to around 40%, involve wastage of a major percentage (up to around 60% of fuel input) to the environment in the form of exhaust gases, as waste heat. Moreover, the gas turbines involve higher exhaust mass flow rates (due to higher specific air flow), higher gas temperature and higher amount of (medium grade) heat in the exhaust than diesel engines **[2]**. In a typical shipboard application, useful shaft power produced by the waste heat recovery can either be fed to a ship's main engine gearbox to increase the ship's speed and thus share the propulsion load with ship's main engines or even drive a generator to produce electrical power and thus share ship's electric load with ship's generators at sea.

The conventional waste heat recovery cycles such as steam and organic Rankine cycles, helium based Brayton cycles, air bottoming cycles and Kalina cycles typically involve large footprint, volume and weight implications rendering them unsuitable for shipboard application due to severe restrictions on size, volume and weight. Supercritical CO_2 cycle (also referred to as 'Feher cycle') is an emerging trend in energy conversion systems because such cycles employ CO_2 (R744) which is a natural, harmless, non-ignitable and chemically stable working fluid/refrigerant having ODP of 0 and GWP of 1.0, as compared to chemical fluids normally used in conventional systems. Supercritical CO_2 power cycles can avoid most of the problems of the conventional cycles and yet retain many of their advantages **[3]**. Several researchers **[3-11]** have reported various features of supercritical CO_2 based cycle, over conventional cycles, favourable for energy sources at medium temperature (400 – 550 ^{0}C) as well as beyond (650 – 800 ^{0}C). Some of these features are (a) more energy-density (b) higher thermal efficiency, (c) least size and weight of the components (d) low volume-to-power ratio (d) least cost (e) no blade-erosion in the turbines (f) no cavitation in the pump (g) single stage turbo-machinery (h) single phase fluid in the heat rejection process and (j) higher thermal efficiency. In addition to these benefits, CO_2 has relatively

lower critical temperature and pressure, as compared to conventional fluids, which makes it relatively easier for CO_2 to attain supercritical state, where it exhibits certain thermos-physical properties that ensure low compression work and efficient heat exchange process.

The research on supercritical CO_2 power cycles has currently focussed on advanced generation nuclear reactors **[8-13]**, concentrated solar power **[14-18]**, automotive **[19]** and various geothermal power application involving high to medium grade energy sources. Dostal **[8-9]** reported a thermal efficiency of 41% at the turbine inlet temperature of 550 ^{0}C for the recuperated power cycle. Ishiyama **[10]** reported a cycle thermal efficiency of 36.4% for a supercritical CO_2 power cycle with a heat source temperature of 480 ^{0}C for a prototype fusion power reactor. Zhang et al. investigated solar CO_2 Rankine systems utilising solar energy as heat source and reported that supercritical CO_2 can be a good working fluid to collect heat efficiently **[14]**. In addition**,** supercritical CO_2 cycles are also being researched in the field of automotive applications. Chen **[19]** investigated supercritical carbon dioxide power cycle for automobiles with engine exhaust as heat source to reduce vehicle fuel's consumption and reported a thermal efficiency of about 19%. The cycle is reported to be highly regenerative, which makes the internal heat exchanger (or regenerator) very important if power production is sought. Bella **[20-21]** investigated a sc-CO_2 based heat recovery cycle in combination with thermoelectric generation to recover waste heat from gas turbine.

The research so far on supercritical CO_2 Brayton (power) systems is primarily based on the first law or the *energy* analysis, which is a traditional method of thermodynamic analysis wherein an energy balance is applied on the system under consideration. However, *energy* analysis provides no information about internal losses and changes in energy quality during a process. The second law analysis overcomes the short coming of the first law and is based on the concept of exergy defined as "the maximum useful work that can be obtained in a process in which the system comes to equilibrium with the surroundings or attains the dead state" **[22]**. More specifically, the exergy analysis is concerned with the degradation of energy during a process, the entropy generation, and the lost-opportunities-to-do-work **[23]**. The second law of thermodynamics has proved to be a very powerful tool in the optimization of complex thermodynamic systems and is required to establish the difference in "quality" between mechanical and thermal energy **[24]**. Second law analysis provides an effective technique for measuring and optimizing performance of a thermal system by accounting for the energy quality and for thermal systems **[25]**.

Sarkar **[26]** undertook an exergetic analysis of a sc-CO_2 recompression cycle with nuclear reactor as heat source. Butcher & Reddy **[27]** performed second law analysis of a steam Rankine cycle based waste heat recovery power generation system. The research so far has primarily focussed on nuclear, solar and automotive based applications of SC-CO_2 cycles. Therefore, in this paper, the authors have proposed a novel waste heat recovery system (WHRS) for augmentation of power for a shipboard application. Two promising sc-CO_2 bottoming cycle variants have been examined. These are regenerative (closed) Brayton cycle (RBC) and regenerative recompression (closed) Brayton cycle (RRCBC). In both these variants, the low and high pressures and temperatures of the cycle are maintained above the critical point of CO_2. Hence, both these variants are termed as supercritical CO_2 or sc-CO_2 cycles.

2. Thermodynamic Analysis

2.1. Shipboard marine gas turbine – the topping cycle.

The broad specifications **[28]** and typical operating conditions of a shipboard marine gas turbine engine are tabulated in **Table 1**.

Table 1 Specifications of a typical marine gas turbine and shipboard operating conditions. Ref. [**28**].

Term	*ISO* condition	Shipboard scenario
Power Output (MW)	25.06	20.58
Power Level (%)	100%	100%
Thermal Efficiency (%)	38	35
Exhaust Mass Flow Rate (kg/sec)	70.5	63.6
Gas Temperature	839.15 K (566 °C)	845.15 K(572 °C)
Gas Pressure(MPa)	-	0.1084
Intake Pressure Loss(Inch, WG)	Nil	8.5
Uptake Pressure Loss(Inch, WG)	Nil	24
Ambient Conditions	298.15 K (15 °C), 1 bar, 60% RH	308.15 K (35°C), 1 bar, 60% RH

2.2. *sc-CO_2 Regenerative Brayton cycle (RBC).*

Fig. 1 shows the schematic layout of shipboard WHRS with RBC. The cycle consists of five principle components: Waste heat-recovery-heat-exchanger (HRHE) where the hot exhaust gases enter and transfer the heat to the working fluid (sc-CO_2) to raise the enthalpy (heat addition process: *6-1* & g_{in}-g_{out}); turbine where mechanical work is extracted from the system by expanding working fluid having high enthalpy (expansion process: *1-2*); an internal heat exchanger called as regenerator or recuperator (process *2-3* & *5-6*) where the enthalpy of turbine left after expansion in turbine, is internally utilized (regeneration process) to raise the enthalpy of the incoming stream of the high pressure working fluid from the compressor, a water cooled pre-cooler where the heat is rejected by the working fluid (heat rejection process: *3-4* & w_{in}-w_{out}) and a compressor where the low pressure working fluid is compressed(compression process: *4-5*) to cycle high pressure and returned back to the HRHE via a regenerator. Then the cycle starts over again and repeats in a closed loop to produce power.

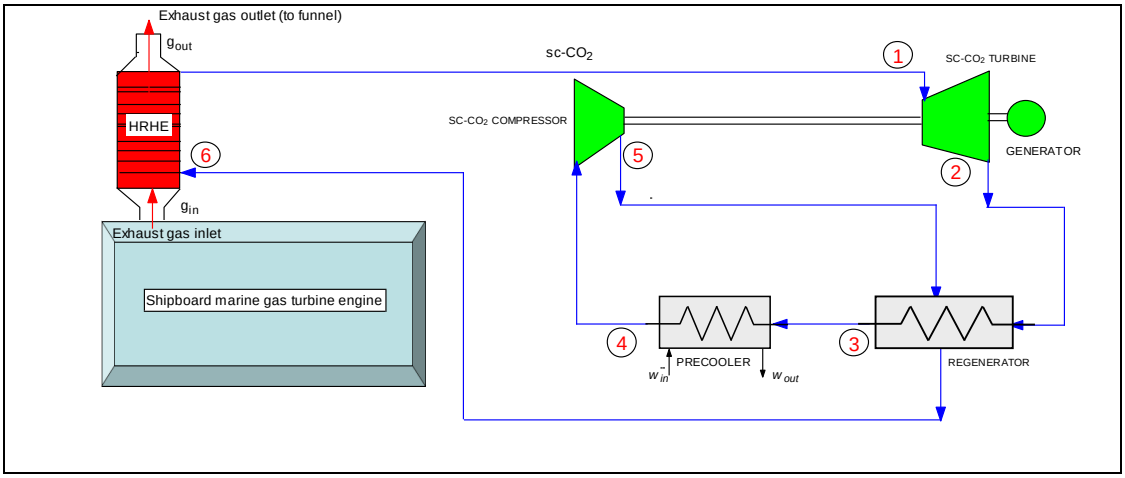


Fig. 1. Schematic layout of shipboard WHRS with RBC.

2.3. *sc-CO_2 Regenerative and Recompression Brayton cycle (RRCBC).*

Fig. 2 shows the schematic layout of RRCBC indicating the components of RRCBC**.** This cycle is quite similar to RBC except for a few changes. In RRCBC, an additional compressor called 're-compression compressor' is introduced. In place of single regenerator in RBC, there are two regenerators: high temperature regenerator (*HTR*) and low temperature regenerator (*LTR*) in RRCBC. Increasing the number of regenerator facilitates effective internal heat exchange with high as well as low temperature working fluid streams in the cycle. In addition, in RRCBC, the mass flow of working fluid at state point ***4*** is split into two streams. The first stream comprising of a fraction (***x***) of the total mass flow, called 're-compressed mass fraction' is routed through the re-compressor and then fed directly to HTR, bypassing the pre-cooler. The second stream having remaining (1-*x*) fraction of the mass flow is allowed to pass through the pre-cooler first, as in RBC, followed by compression in main-compressor. Both streams are merged (via a flow-merger) at state (***7***), prior entering HTR.

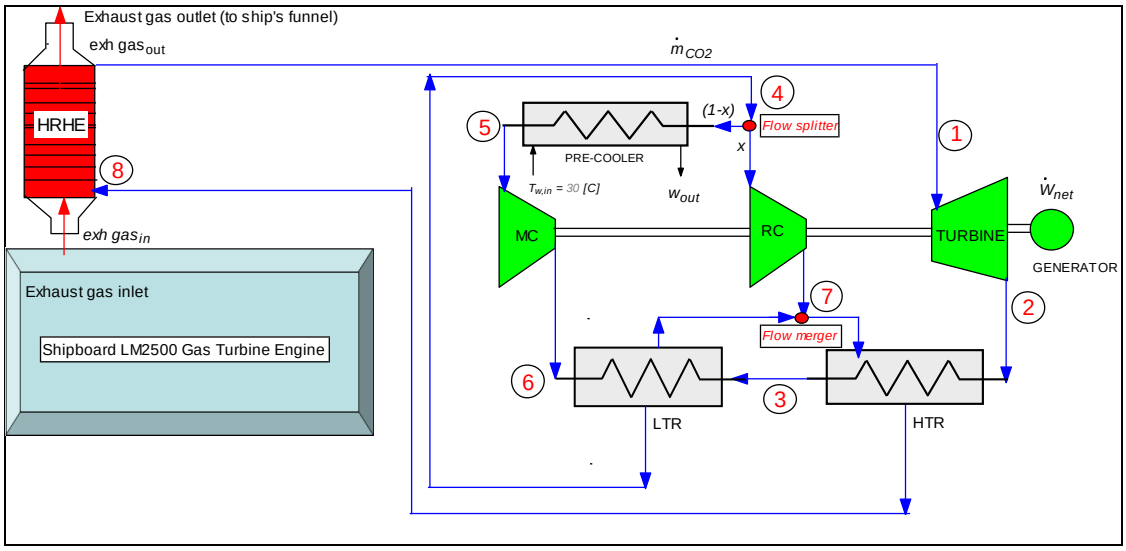


Fig. 2. Schematic layout of shipboard WHRS with RRCBC.

2.4. Assumptions

The following assumptions have been made to simplify the mathematical modelling:

(1) A steady-state steady-flow (*SSSF*) and closed loop operation is considered.
(2) Each component in the cycle is considered to have adiabatic boundaries and be a single control volume.
(3) The kinetic and potential energy terms and their exergy effects are neglected.
(4) The exhaust gas outlet temperature (after waste heat recovery in HRHE) is maintained above its saturation temperature of 549.15 K (276 ^{0}C).
(5) The exergy carried out by the outgoing streams: gas outlet from HRHE to funnel and cooling water outlet stream from pre-cooler to ambient is not considered since currently not being utilized further.

2. 5. Thermodynamic model

Both cycles are thermodynamically modelled based on the mass, energy and exergy balance of individual components and considering the above assumptions. The salient equations and assumptions thereof for energy and exergy analysis are shown in **Table 2**. The relative non-dimensional pressure drop fractions, $\delta\overline{P} = \Delta P / P_i$ have been defined to signify the pressure drop across the working fluid streams per unit inlet pressure (P_i) across the heat exchangers of both cycles.

Table 2 Salient thermodynamic equations along with assumptions for energy and exergy analysis

Component	RBC	RRCBC
Turbine	$\eta_{isen,turbine,RBC} = (h_1 - h_2)/(h_1 - h_{2s})$ $\dot{I}_{turbine} = \dot{m}_{CO_2}(e_1 - e_2) - W_{turbine}$	*Same as in RBC*
Compressor	$\eta_{isen,compressor,RBC} = (h_{5s} - h_4)/(h_5 - h_4)$ $\dot{I}_{compressor,RBC} = \dot{m}_{CO_2}(e_4 - e_5) + W_{compressor}$	***MC*:** $\eta_{isen,mc,RRCBC} = (h_{6s} - h_5)/(h_6 - h_5)$ $\dot{I}_{mc} = \dot{m}_{CO_2}(1-x)(e_5 - e_6) + W_{mc}$ ***RC*:** $\eta_{isen,rc,RRCBC} = (h_{7s} - h_4)/(h_7 - h_4)$ $\dot{I}_{rc} = \dot{m}_{CO_2} x(e_4 - e_7) + W_{rc}$
HRHE	$\varepsilon_{HRHE,RBC} = (T_{g,in} - T_{g,out})/(T_{g,in} - T_6)$ here we have assumed that the heat capacity (the product of isobaric	$\varepsilon_{HRHE,RRCBC} = (T_{g,in} - T_{g,out})/(T_{g,in} - T_8)$ here we have assumed that the heat capacity of the hot stream (*gas side*) is the *minimum.*

	specific heat and the mass flow rate) of the hot stream (*gas*) is the minimum. $\dot{m}_{CO_2}\left(h_1-h_6\right)=m_g\left(h_{g,in}-h_{g,out}\right)$ ***(EBE)*** $\dot{I}_{HRHE}=\dot{E}_{in,RBC}+\dot{m}_{CO_2}\left(e_6-e_1\right)$	$\dot{m}_{CO_2}\left(h_1-h_8\right)=m_g\left(C_{p,g,in}T_{g,in}-C_{p,g,out}T_{g,out}\right)$ ***(EBE)*** $\dot{I}_{HRHE}=\dot{E}_{in,RRCBC}+\dot{m}_{CO_2}\left(e_8-e_1\right)$
Regenerator	$\varepsilon_{regenerator,RBC}=(h_2-h_3)/(h_2-h_{5'})$, here we have assumed that the minimum heat capacity is in the hot stream (2-3) of the regenerator and $h_{5'}$ represents the enthalpy of sc-CO_2 at T=T_5 and P=P_3. $h_2-h_3=h_6-h_5$ ***(EBE)*** $\dot{I}_{regenerator,RBC}=\dot{m}_{CO_2}\left\{\left(e_2-e_3\right)+\left(e_5-e_6\right)\right\}$	***HTR*:** $\varepsilon_{HTR}=(h_2-h_3)/(h_2-h_{7'})$, here we have assumed that the minimum heat capacity is in the hot stream (2-3) of the HTR and $h_{7'}$ represents the enthalpy of sc-CO_2 at T=T_7 and P=P_3. $h_2-h_3=h_8-h_7$ ***(EBE)*** $\dot{I}_{HTR}=\dot{m}_{CO_2}\left\{\left(e_2-e_3\right)+\left(e_7-e_8\right)\right\}$ ***LTR*:** $\varepsilon_{LTR}=\left(h_3-h_4\right)/\left(h_3-h_{6'}\right)$, here we have assumed that the minimum heat capacity is in the hot stream (3-4) of LTR and $h_{6'}$ represents the enthalpy of sc-CO_2 at T=T_6 and P=P_4. $h_3-h_4=\left(1-x\right)(h_7-h_6)$ ***(EBE)*** $\dot{I}_{LTR}=\dot{m}_{CO_2}\left\{\left(e_3-e_4\right)+\left(1-x\right)\left(e_6-e_7\right)\right\}$
Precooler	$\dot{m}_{CO_2}\left(h_3-h_4\right)=m_w\left(h_{w,out}-h_{w,in}\right)$ ***(EBE)*** $\dot{I}_{precooler,RBC}=\dot{m}_{CO_2}\left(e_3-e_4\right)$	$\left(1-x\right)\dot{m}_{CO_2}\left(h_4-h_5\right)=m_w\left(h_{w,out}-h_{w,in}\right)$ ***(EBE)*** $\dot{I}_{precooler,RRCBC}=\dot{m}_{CO_2}\left(1-x\right)\left(e_4-e_5\right)$

2.5.1. RBC

The rate of heat transferred by the waste heat source (*gas*) to the working fluid (*sc-CO_2*) in the HRHE of RBC is given by

$$\dot{Q}_{in,RBC}=\dot{m}_{CO_2}\left(h_1-h_6\right) \quad (1)$$

The rate of heat rejected by the working fluid (*sc-CO_2*) in pre-cooler of RBC is given by

$$\dot{Q}_{out,RBC}=\dot{m}_{CO_2}\left(h_3-h_4\right) \quad (2)$$

The actual expansion work produced in turbine is given by

$$\dot{W}_{turbine,RBC}=\dot{m}_{CO_2}\left(h_1-h_2\right) \quad (3)$$

The expression for actual compression work input in compressor is given by

$$\dot{W}_{compressor,RBC}=\dot{m}_{CO_2}\left(h_5-h_4\right) \quad (4)$$

The actual net work produced by the RBC is given by

$$\dot{W}_{net,RBC}=\dot{W}_{turbine,RBC}-\dot{W}_{compressor,RBC} \quad (5)$$

The first law efficiency (also called as the energetic or the thermal efficiency) of the RBC is given by

$$\eta_{I,RBC(\%)}=\left(\dot{W}_{turbine,RBC}-\dot{W}_{compressor,RBC}\right)/\dot{Q}_{in,RBC} \quad (6)$$

The *Back-Work ratio* (*BWR*) of a thermodynamic cycle represents the ratio of the work consumed by the compressor(s) and the work generated by the turbine(s). For RBC, it is given by:

$$BWR_{RBC}=\left(\dot{W}_{compressor,RBC}/\dot{W}_{turbine,RBC}\right) \quad (7)$$

For the second law analysis, the specific exergy flow at each state point (i) is determined from the following general expression for SSSF:

$$e_i = \left(h_i - h_0\right) - T_0\left(s_i - s_0\right) \qquad \text{where } T_0 = 308.15 \text{ K}, P_0 = 0.101325 \text{ MPa; For RBC: } i = 1 \text{ to } 6 \tag{8}$$

The total exergy input rate of RBC, assuming constant heat source temperature (gas inlet temperature), $T_{g,in}$ (K), is expressed as:

$$\dot{E}_{in,RBC} = \dot{Q}_{in,RBC}\left(1 - T_0 / T_{g,in}\right) \tag{9}$$

Based on component irreversibility rate (expressions given in **Table 2**), the total irreversibility rate of RBC is:

$$\dot{I}_{Total,RBC} = \dot{I}_{turbine} + \dot{I}_{compressor} + \dot{I}_{regenerator} + I_{precooler} + I_{HRHE} \tag{10}$$

The non-dimensional fraction (i_j) of irreversibility for the j^{th} component of a cycle can be defined as the ratio of the irreversibility rate of the j^{th} component and the total exergy input rate of the cycle **[26]**. Thus,

$$i_j = \dot{I}_j / \dot{E}_{in,RBC} \tag{11}$$

The second law efficiency or the exergy efficiency of the *RBC* can be expressed as the ratio of net exergy output to the exergy input through HRHE. Alternatively, the exergy efficiency can also be expressed in terms of the sum of exergy destruction occurring in each component of the cycle. Thus,

$$\eta_{II,RBC} = \dot{W}_{net,RBC} / \dot{E}_{in,RBC} = 1 - \sum_{j=1}^{6} i_j \tag{12}$$

2.5.2. RRCBC

For RRCBC, the rate of heat transferred from the waste heat source (gas) to the working fluid (CO_2) in the HRHE is given by

$$\dot{Q}_{in,RRCBC} = \dot{m}_{CO_2}\left(h_1 - h_8\right) \tag{13}$$

The rate of heat rejected by the working fluid in the precooler is given by

$$\dot{Q}_{out,RRCBC} = \left(1 - x\right)\dot{m}_{CO_2}\left(h_4 - h_5\right) \tag{14}$$

The actual expansion work produced in turbine is given by

$$\dot{W}_{turbine} = \dot{m}_{CO_2}\left(h_1 - h_2\right) \tag{15}$$

The expression for actual compression work input in main compressor is given by

$$\dot{W}_{mc} = \left(1 - x\right)\dot{m}_{CO_2}\left(h_6 - h_5\right) \tag{16}$$

The expression for actual compression work input in re-compressor is given by

$$\dot{W}_{rc} = x\dot{m}_{CO_2}\left(h_6 - h_5\right) \tag{17}$$

The actual net work produced, the first law efficiency and BWR of RRCBC are respectively expressed as

$$\dot{W}_{net} = \dot{W}_{turbine} - \left(\dot{W}_{mc} + W_{rc}\right) \tag{18}$$

$$\eta_{I,RRCBC} = \dot{W}_{net} / \dot{Q}_{in,RRCBC} \tag{19}$$

$$BWR_{RRCBC} = \left(\dot{W}_{mc} + \dot{W}_{rc}\right) / W_{turbine} \tag{20}$$

The total exergy input in the RRCBC is given by:

$$\dot{E}_{in,RRCBC} = \dot{Q}_{in,RRCBC}\left(1 - T_0 / T_{g,\,in}\right) \tag{21}$$

Based on component irreversibility rate (expressions given in **Table 2**), the total irreversibility rate of the RRCBC is:

$$\dot{I}_{Total,RRCBC} = \dot{I}_{turbine} + \dot{I}_{mc} + \dot{I}_{rc} + \dot{I}_{LTR} + I_{HTR} + I_{precooler} + I_{HRHE} \tag{22}$$

The second law efficiency (or the exergetic efficiency) for the RRCBC is expressed as

$$\eta_{II,RRCBC} = \dot{W}_{net,RRCBC} / \dot{E}_{in,RRCBC} = 1 - \sum_{j=1}^{8} i_{j,RRCBC} \quad (23)$$

The various non-dimensional fractions (i_j) of irreversibility corresponding to the j^{th} component in RRCBC are defined in a manner similar to that for RBC.

2.6. Simulation

The thermodynamic analyses have been performed with EES (V8.158) **[29]**. For thermodynamic properties of CO_2, the EES in-built library was used. It provides high accuracy thermodynamic properties for CO_2 using the fundamental equation of state developed in the literature **[30]**. A simulation code has been developed for performance evaluation of both variants of sc-CO_2 power cycle.

3. Results and Discussion

The optimum cycle-pressure-ratio corresponding to the maximum cycle efficiency was determined for various operating and design parameters with the variation of the cycle pressure ratio having a step size of 0.001 MPa and range of 2.0 to 2.7. The upper bound of 2.7 was selected in order to ensure the supercritical phase of the working fluid at the minimum cycle pressure (compressor inlet pressure) through restricting the minimum cycle pressure *above* the critical point pressure (7.37 MPa). The maximum cycle pressure was fixed at 20 MPa.

3.1. Input parameters

Unless otherwise specified, the *standard operating & design conditions* shall form the input parameters presented in **Table 3** while evaluating and discussing the performance of RBC and RRCBC.

Table 3 Standard operating & design conditions for sc-CO_2 waste heat recovery cycle.

Operating and design data	**Set value**
Shipboard marine gas turbine engine load (%)	100
Gas mass flow rate (kg/s)	63.6
Gas inlet temperature (K)	845.15 (572 ^oC)
Gas inlet pressure (MPa)	0.1084
sc-CO_2 turbine inlet or maximum cycle temperature (K)	823.15 (550 ^oC)
sc-CO_2compressor inlet or minimum cycle temperature (K)	313.15 (40 ^oC)
sc-CO_2compressor discharge or maximum cycle pressure(MPa)	20
sc-CO_2compressor inlet or minimum cycle pressure (MPa)	>7.37 MPa
sc-CO_2 pressure ratio(-)	2.5 or as specified
PPTD in HRHE (hot side)	22
PPTD in pre-cooler (hot side)	15
Cooling water inlet temperature (K)	303.15 (30 ^oC)
Component pressure loss in heat exchangers (%)	Nil
Turbine isentropic efficiency (%)	90
Compressor isentropic efficiency (%)	85
Effectiveness of regenerator (%)	86
Effectiveness of HTR (%)	86
Effectiveness of LTR (%)	86
Effectiveness of HRHE (%)	90

3.2. Validation of the model

The mathematical model has been verified with the literature data **[9]**. The comparison of performance results for the basic design with *conservative* efficiencies of turbo-machinery, maximum cycle temperature of 550 ^{0}C, minimum cycle temperature of 32 ^{0}C, maximum cycle pressure of 20 MPa, main compressor isentropic efficiency of 89%, re-compressor isentropic efficiency of 89% and turbine isentropic efficiency of 90%, HTR effectiveness of 96.3%, LTR effectiveness of 92.1% and cycle pressure ratio of 2.6 (*optimum* value) shows that the results obtained for the optimised RRCBC: recompression mass fraction (0.38) and cycle energetic efficiency (46.56%) values are closely matching with literature results (0.40 and 45.27% respectively). The validation of the RRCBC model also implies validation of the RBC model since both these models are identical, while RBC is a special case of RRCBC, where the re-compressed mass fraction approaches zero value.

3.3. Temperature-Entropy (T-s) diagram and Exergy Balance Sheet

Figs. 3(a) & **3(b)** show the *T-s* diagrams of RBC and RRCBC respectively at *optimum* cycle pressure ratios.

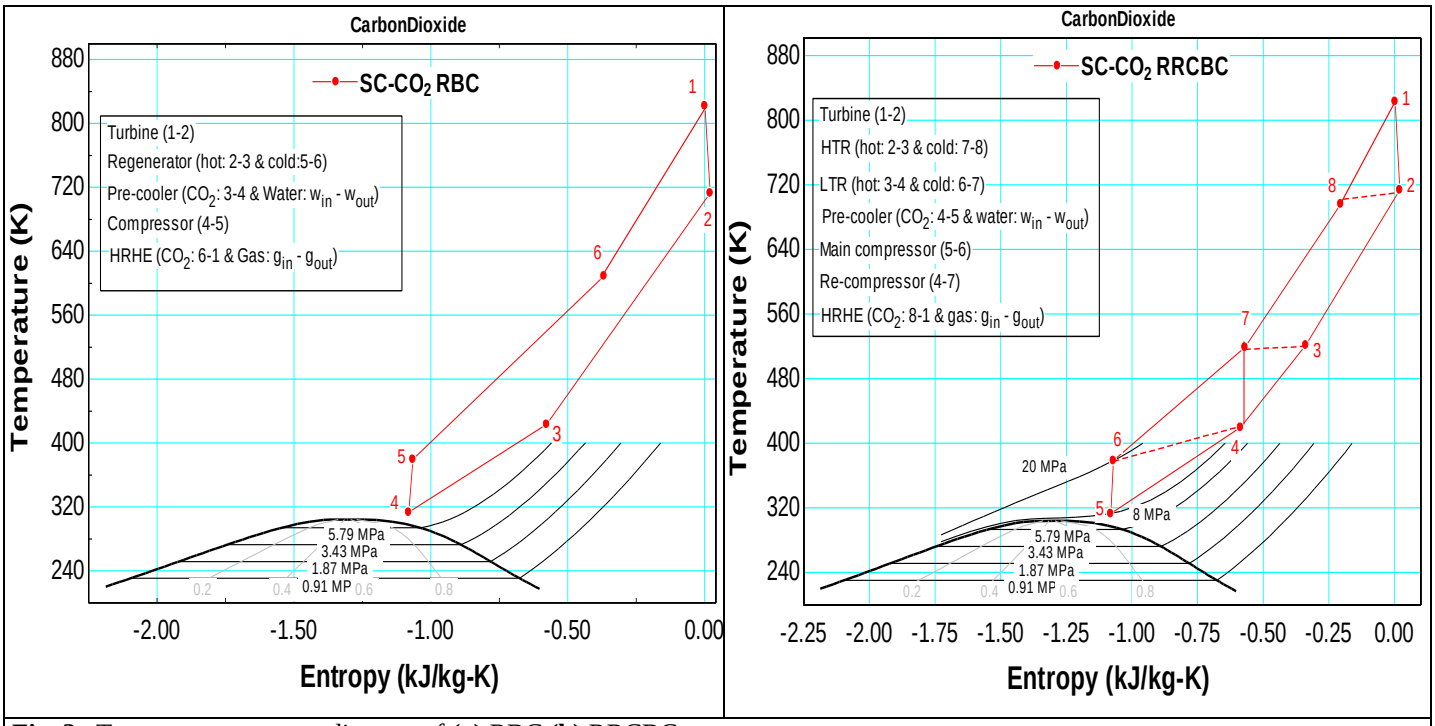


Fig. 3. Temperature-entropy diagram of **(a)** RBC **(b)** RRCBC.

The exergy balance sheets for RBC and RRCBC are presented in **Tables 4 & 5** respectively. Results show that, in both variants, the irreversibility of the heat exchangers are significantly more compared to that of turbomachinery and the HRHE and pre-cooler irreversibility are most significant due to higher heat exchange temperature difference. The regenerator and pre-cooler irreversibility is relatively more pronounced in RBC as compared to RRCBC.

Table 4 Exergy balance of RBC (Standard operating & design condition, optimum pressure ratio).

Exergy	kJ/kg	%(of exergy input)	Irreversibility	kJ/kg	%(of exergy input)	%(of total irreversibility)
Exergy input to the cycle	167.9	100.0	HRHE	18.3	10.9	22.3
			Precooler	31.9	19.0	39.0
			Regenerator	16.5	9.8	20.2

Exergy recovered through Turbine	152.6		Turbine	7.7	4.6	9.5
Exergy input to Compressor	66.5		Compressor	7.4	4.4	9.0
Net exergy output	86.1	51.3	Total Irreversibility	81.8	48.7	100.0

Table 5 Exergy balance of RRCBC (Standard operating and design condition, optimum pressure ratio).

Exergy	**kJ/kg**	**% (of exergy input)**	**Irreversibility**	**kJ/kg**	**% (of exergy input)**	**% (of total irreversibility)**
Exergy input to the cycle	122.7	100.0	HRHE	9.8	8.0	21.1
			Precooler	12.3	10.1	26.7
			HTR	8.2	6.7	17.7
			LTR	6.7	5.5	14.5
Exergy recovered through Turbine	119.0		Turbine	5.7	4.7	12.4
Exergy input to MC	27.3		MC	2.5	2.0	5.3
Exergy input to RC	15.3		RC	1.1	0.9	2.3
Net exergy output	76.4	62.3	Total Irreversibility	46.3	37.7	100.0

3.4. Performance Results

The *optimised* performance results obtained for RBC & RRCBC are shown in **Table 6**. Results show that the shipboard propulsion plant thermal efficiency enhances from 35% to 44.3% (RBC) and 45.1% (RRCBC) and its net work output increases by 26% (RBC) and 25% (RRCBC) at 100% load. It is also seen that the optimised pressure ratio ($r_{p,opt}$) of the RRCBC is significantly (9.2%) lower than that of the RBC. The parametric analyses of both variants is discussed in detail in the succeeding sub-paragraphs.

Table 6 Optimised performance results of RBC & RRCBC.

Cycle	$r_{p,opt}$ (*)	$\eta_{I,cycle}$ (%)	$\eta_{II,cycle}$ (%)	$\eta_{I,CC}$ (%)	$\eta_{I,WHRS}$ (%)	$\eta_{II,WHRS}$ (%)	$\dot{W}_{net}$ (kW)	$\dot{m}_{CO_2}$ (kg/s)
RBC	2.7	32.1	50.5	44.3	9.1	34.5	5402	64.7
RRCBC	2.5	39.6	62.3	45.1	8.6	32.9	5148	67.4

Note (*): $r_{p,opt}$ corresponds to the *maximum* bottoming cycle efficiency at the *standard* operating and design condition.

Effect of cycle pressure ratio

Fig. 4 (a) & **4(b)** illustrate the influence of cycle-pressure-ratio on energetic and exergetic efficiency by varying the pressure ratio from 2.0 up to 8.0 with all other parameters fixed. It is clearly seen that both cycle variants are very sensitive to cycle pressure ratio and that even a small departure from the optimum pressure ratio

can cause a significant reduction in cycle efficiency. This trend, in both cycle variants is on account of two factors. Firstly, for a fixed heat exchanger volume as the pressure ratio is varied, the heat exchangers (including HTR and LTR in RRCBC) effectiveness (and thus cycle efficiency) is affected because of the variations in the working fluid density which affects the velocity and pressure drops. Thus, at high pressure ratios, the density of the working fluid at the high pressure side of heat exchangers becomes lower resulting in higher velocity and higher pressure drops. This steep increase in pressure drops with the pressure ratio at high pressure side of the heat exchangers (including LTR and HTR) contributes significantly to the deterioration of the cycle efficiency at high pressure ratios. For a fixed heat exchanger volume, at low pressure ratios, it is the HTR that is responsible for the significant reduction in cycle efficiency and at high pressure ratios; the LTR also contributes albeit lower than HTR's contribution. Secondly, the heat available for regeneration also decreases as the pressure ratio increases which also contributes in reduction in overall cycle efficiency **[9]**. Also, it is seen that the RRCBC efficiency reduction due to deviation from optimum pressure ratio is relatively more pronounced. In case of the recompression cycle, there are two compressors working in parallel and there is a *flow-split* i.e. certain mass fraction (recompressed mass fraction) of the working fluid is bypassed from sending to the pre-cooler and instead it is routed to HTR via re-compression compressor. Due to flow split, in RRCBC, when the pressure ratio (the main compressor pressure ratio) is varied, the *flow-split* changes as well which causes the cycle efficiency to drop quickly if the cycle operates away from its optimum pressure ratio **[8, 9]**. Due to this, RRCBC efficiency is more sensitive to the cycle-pressure-ratio as compared to RBC.

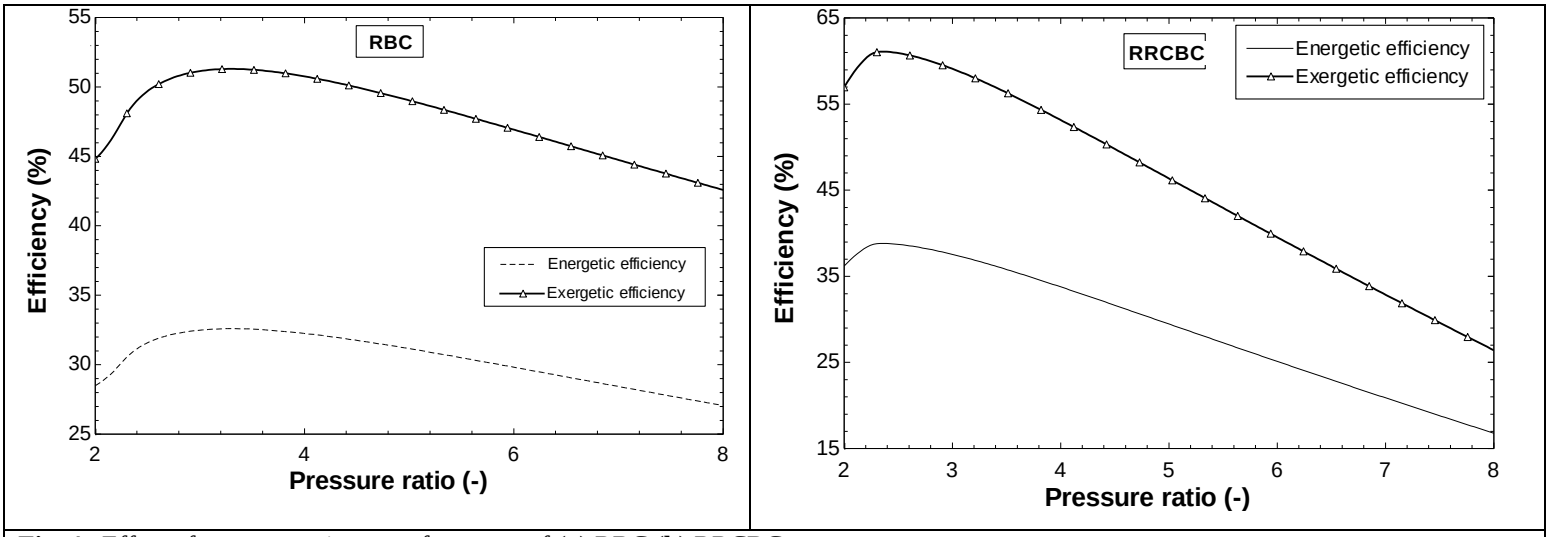


Fig. 4. Effect of pressure ratio on performance of **(a)** RBC **(b)** RRCBC.

Effect of minimum cycle temperature

Figs. 5 (a) & **5(b)** depict the dependence of the *optimum* cycle pressure ratio and cycle efficiency on the minimum cycle temperature. It is observed that for both the cycle variants, the cycle efficiency decreases almost linearly with increase in minimum cycle temperature. The efficiency is high at low temperatures owing to higher difference between maximum and minimum cycle temperature (temperature lift). This lift decreases linearly with increase in minimum cycle temperature resulting in linear decrease in cycle efficiency. In both cycle variants, an increase in the minimum cycle temperature causes almost a linear drop in optimum pressure ratio value. However, in case of RRCBC, at some point, optimum pressure ratio value saturates when the compressor operates far enough away from the critical point and the fluid properties are no longer significantly affected by it. The increase in the minimum cycle temperature from 32 to 50 ^{o}C causes a reduction in the efficiency of RBC and RRCBC by about 7% and about 11% respectively, which indicates that the effect of minimum cycle temperature is more significant (1.6 times) for RRCBC than RBC.

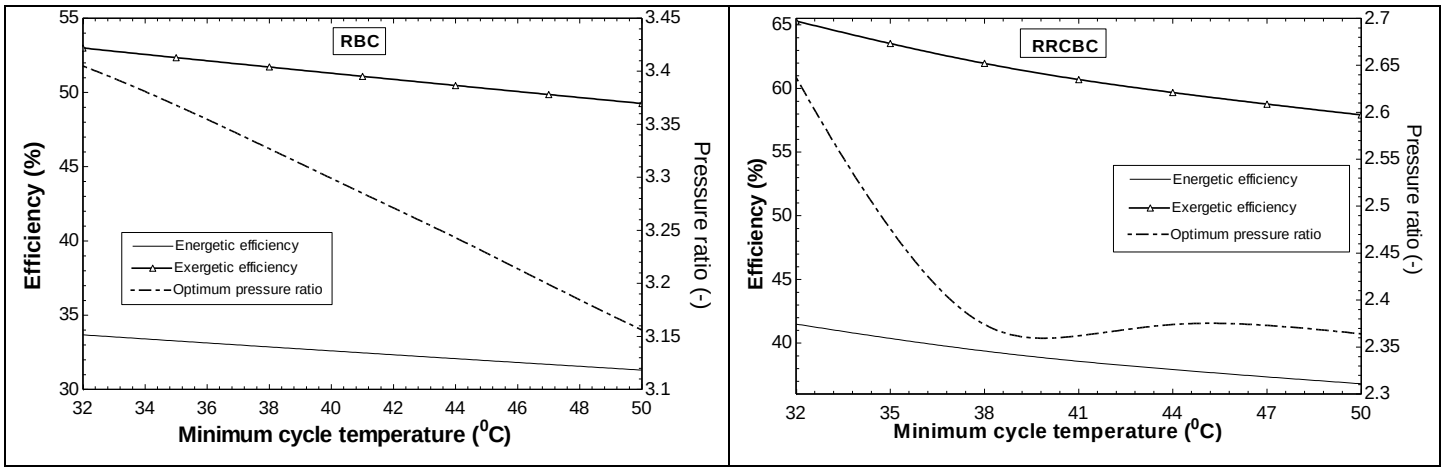


Fig. 5. Effect of minimum cycle temperature on cycle performance and optimum pressure ratio for (a) RBC (b) RRCBC.

Effect of maximum cycle temperature

Figs. 6 (a) & **6(b)** illustrate the effect of maximum operating (turbine inlet) temperature on the optimum cycle pressure ratio and corresponding energetic and exergetic efficiencies. As the maximum cycle temperature increases, the first and the second law efficiency of both cycles also increases almost linearly. This is expected since increasing the turbine inlet temperature increase the temperature lift and thus must improve the underlying thermodynamic efficiency of a Brayton cycle. The effect of turbine inlet temperature on optimum compressor pressure ratio is negligible compared to compressor inlet temperature since the specific heat capacity values of CO_2 at all state points are nearly invariant with the increase in turbine inlet temperature, as it is far away from the critical point unlike to the compressor inlet temperature. Due to this reason, the effect of minimum operating temperature on the cycle performance is much more predominant than the maximum operating temperature in both cycle variants. Further, the increase in maximum operating temperature from 450 to 550 ^{o}C, causes an increase in efficiency of RBC by almost 13 % and of RRCBC by almost 10%, indicating that the effect of maximum cycle temperature in case of RBC is more significant (1.3 times) than RRCBC.

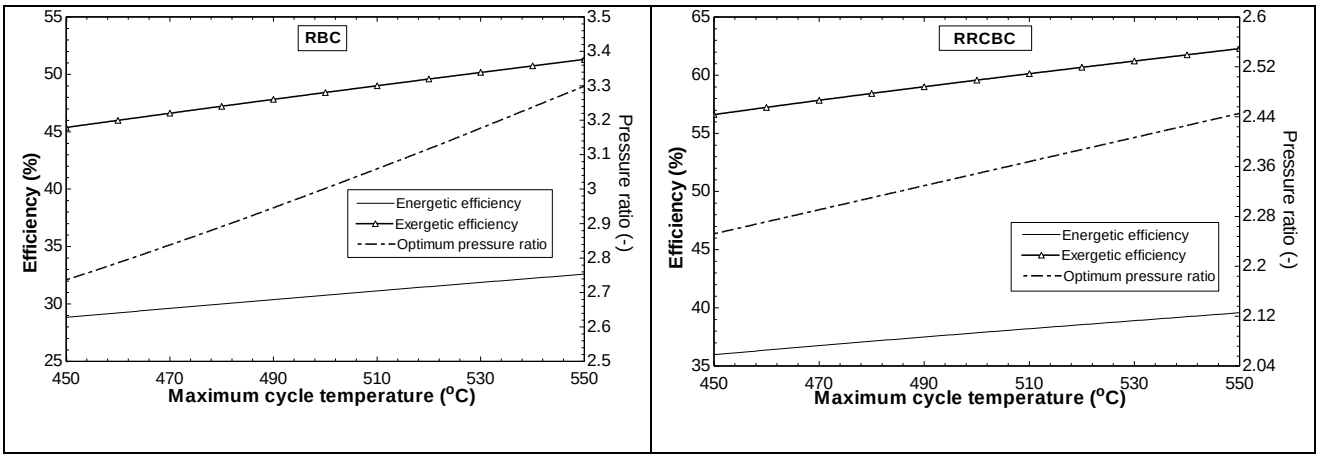


Fig. 6. Effect of maximum cycle temperature on cycle efficiency and optimum pressure ratio in (a) RBC (b) RRCBC.

Effect of component pressure drop

The effect of the fractional pressure drops (varied from 0 to 3 %) in the heat exchangers on exergetic efficiency of bottoming cycle variants is presented in **Figs. 7 (a)** & **7(b)**. It can be clearly seen that that the efficiency decreases linearly with the increase in component pressure drop. This is logical since the only effect of heat exchanger pressure drops in a cycle is its reduction of the turbine pressure ratio which is directly proportional to the turbine work. In RBC, the effect of pressure drop in HRHE and regenerator is more significant than other components, whereas the effect of pre-cooler pressure drop is least significant on the second law efficiency reduction. In RRCBC too, the effect of pressure drop in HRHE is most significant. The next most significant component after HRHE is the HTR and then LTR, whereas the effect of pre-cooler pressure drop (as in RBC) is

the least significant on the second law efficiency reduction. In RRCBC, in case of LTR, the effect of the cold stream pressure drop is less significant as compared to the hot stream pressure drop due to a lesser mass flow rate (only (1-*x*) fraction) in the cold-side of LTR. In general, pressure drop of hot streams of regenerator & HRHE, for RBC and hot streams of HTR, LTR & HRHE, for RRCBC exhibit significantly greater effect on reduction in cycle efficiency because of lower fluid density in these regions.

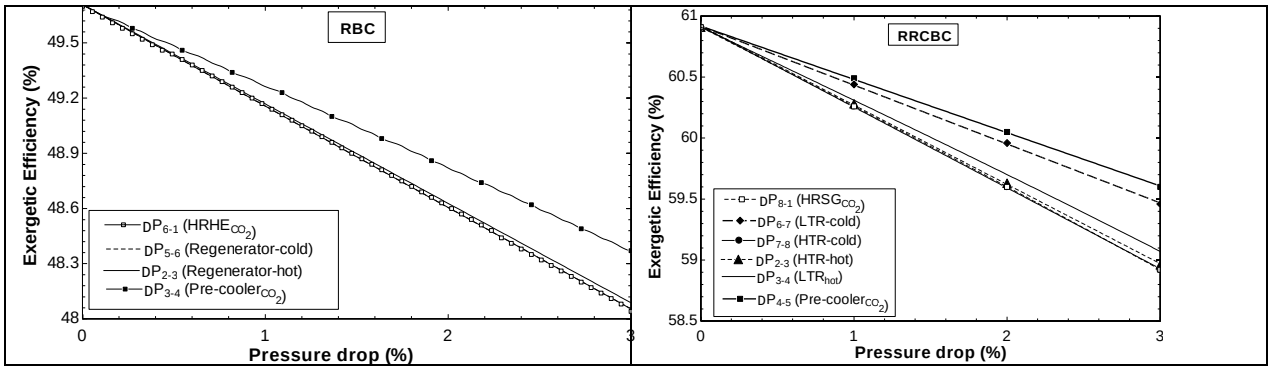


Fig. 7. Effect of component pressure drop on performance of (a) RBC (b) RRCBC.

4. Conclusions

The energetic and exergetic performances analyses and optimisation of two most promising variants of sc-CO_2 Brayton cycles for waste heat recovery from shipboard gas turbine engine has been carried out. The salient conclusions are as follows:

- Shipboard plant thermal efficiency enhances from 35% to 44.3% (RBC) and 45.1% (RRCBC) and its net work output increases by 26% (RBC) and 25% (RRCBC), at 100% load.
- An optimum pressure ratio exists for the bottoming cycle that maximises energetic as well as exergetic efficiency. RRCBC efficiency is more sensitive to the cycle-pressure-ratio which implies that any deviations from the optimum pressure ratio will lead to deterioration of performance, more significant in RRCBC variant.
- In both cycles, the effect of irreversibilities of heat exchangers is significantly more on exergetic performance, as compared to that of turbomachinery.
- The regenerator and pre-cooler irreversibility is relatively more pronounced in RBC compared to RRCBC.
- The effect of minimum cycle temperature is more significant (about 1.6 times) for RRCBC than RBC.
- The effect of the minimum operating temperature on the optimum cycle pressure ratio and cycle performance is more pronounced than the maximum operating temperature.
- The effect of pressure drops in HRHE & HTR, on deterioration of the second law efficiency, is more pronounced than other components. The effect of precooler pressure drop is the least on the second law efficiency reduction.

References

[1] Executive summary, International Energy Agency (IEA), World Energy Outlook, WEO2015, available on Internet website https://www.iea.org.

[2] R.L. Harrington, Marine Engineering, Society of Naval Architects and Marine Engineers (1992).

[3] E.G. Feher, The supercritical thermodynamic power cycle, presented to the IECEC, Florida, Douglas paper no. 4348 (1967).

[4] E.G. Feher, The supercritical thermodynamic power cycle, Int. J. Energy Conversion 8 (1968) 85-90.

[5] G. Angelino, Real gas effects in carbon dioxide cycles, American Society of Mechanical Engineers, Gas Turbine Conference and Products Show V001T01A071 (1969).

[6] O.V. Combs, An investigation of the supercritical CO_2 cycle (Feher cycle) for shipboard application, Doctoral dissertation, Massachusetts Institute of Technology (1977).

[7] H. D. Marron, Gas turbine waste heat recovery propulsion for U.S. Navy surface combatants, Naval Engineers Journal 93(5) (1981) 65–75.

[8] V. Dostal, M.J. Driscoll, P. Hejzlar, A supercritical carbon dioxide cycle for next generation nuclear reactors. Massachusetts Institute of Technology, Dept. of Nuclear Engineering, Cambridge, MA, (2004) Paper No. MIT-ANP-TR-100.

[9] V. Dostal, A supercritical carbon dioxide cycle for next generation nuclear reactors, Ph.D. thesis, MIT, USA (2004).

[10] S. Ishiyama, Y. Muto, Y. Kato, S. Nishio, T. Hayashi, Y. Nomoto, Study of steam, helium and supercritical CO2 turbine power generations in prototype fusion power reactor, Progress in Nuclear Energy 50(2) (2008) 325-332.

[11] H.S. Pham, N. Alpy, J.H. Ferrasse, O. Boutin, J. Quenaut, M. Tothill, D. Haubensack, M. Saez, Mapping of the thermodynamic performance of the supercritical CO_2 cycle and optimization for a small modular reactor and a sodium-cooled fast reactor, Int. J. Energy 87 (2015) 412-424.

[12] E.M. Clementoni, T.L. Cox, M.A. King, Off-Nominal component performance in a supercritical carbon dioxide Brayton cycle, J. Eng. for Gas Turbines and Power 138(1) (2016) 011703.

[13] H. Zhao, Q. Deng, W. Huang, D. Wang, Z. Feng, Thermodynamic and economic analysis and multi-objective optimization of supercritical CO_2 Brayton cycles, J. Eng. for Gas Turbines and Power 138(8) (2016) 081602.

[14] X.R. Zhang, H. Yamaguchi, D. Uneno, K. Fujima, M. Enomoto, N. Sawada, Analysis of a novel solar energy-powered Rankine cycle for combined power and heat generation using supercritical carbon dioxide, J. Renewable Energy 31(12) (2006) 1839-54.

[15] X.R. Zhang, H. Yamaguchi, K. Fujima, M. Enomoto, N. Sawada, Theoretical analysis of a thermodynamic cycle for power and heat production using supercritical carbon dioxide, Energy 32(4) (2007) 591-9.

[16] P. Garg, P. Kumar, K. Srinivasan, Supercritical carbon dioxide Brayton cycle for concentrated solar power, Int. J. Supercritical Fluids 76 (2013) 54-60.

[17] T.W. Neises, C.S. Turchi, Supercritical CO_2 power cycles: design considerations for concentrating solar power, 4th Int. Symposium—Supercritical CO_2 (2014) 9-10.

[18] X. R. Zhang, H. Yamaguchi, D. Uneno, Experimental study on the performance of solar Rankine system using supercritical CO2, Renewable Energy 32(15) (2007) 2617-28.

[19] Y. Chen, P. Lundqvist, P. Platell, Theoretical research of carbon dioxide power cycle application in automobile industry to reduce vehicle's fuel consumption, Appl. Therm. Eng. 25(14) (2005) 2041-2053.

[20] F.A. Di, Bella, Gas turbine engine exhaust waste heat recovery navy shipboard module development, Int. Supercritical CO2 Power Cycle Symposium (2011).

[21] F.A. Di, Bella, Gas turbine engine exhaust waste heat recovery using supercritical CO_2 Brayton cycle with thermoelectric generator technology, InASME, 9th Int. Conf. on energy sustainability collocated with the ASME 2015 power conference, the ASME 2015 13th Int. Conf. on fuel cell science, engineering and technology, and the ASME 2015 Nuclear Forum (2015) pp. (V001T04A003-V001T04A003).

[22] T.J. Kotas, The Exergy Method of Thermal Plant Analysis, Elsevier (2013).

[23] A. Bejan, Fundamentals of exergy analysis, entropy generation minimization, and the generation of flow architecture, Int. J. Energy Research 26(7) (2002) 0-43.

[24] Y.A. Cengel, M.A. Boles, Thermodynamics – An Engineering Approach, Tata McGraw Hill, New Delhi (2002).

[25] I. Dincer, Thermal energy storage systems as a key technology in energy conservation, Int. J. Energy Research 26(7) (2002) 567-88.

[26] J. Sarkar, Second law analysis of supercritical CO2 recompression Brayton cycle, Energy 34(9) (2009) 1172-8.

[27] C.J. Butcher, B.V. Reddy, Second law analysis of a waste heat recovery based power generation system, Int. J. Heat and Mass Transfer 50(11) (2007) 2355-63.

[28] Internet website, https://www.geaviation.com/marine/engines/military/lm2500 regarding specifications of shipboard marine gas turbine engine (Last accessed on 07 Mar 2017) and technical interaction with GE regarding shipboard aeroderivative marine gas turbines.

[29] S.A. Klein, EES, Engineering Equation Solver, version 8.158, F Chart Software, Middleton, WI (2008).

[30] R. Span, W. Wagner, A new equation of state for carbon dioxide covering the fluid region from the triple point temperature to 1100 K at pressures up to 800 MPa, J. Physical and Chemical Reference Data, 25(6) (1996) 1509-96.

Nomenclature		***Subscripts***	
BWR	*back-work-ratio*	CO_2	Carbon Dioxide (working fluid)
C_p	isobaric specific heat, kJ/kg K	*g*	gas
e_j	specific exergy or exergy per unit mass basis at j^{th} state point of the thermodynamic cycle, kJ/kg	*HRHE*	heat recovery heat exchanger
$\dot{E}$	exergy rate or exergy per unit time basis, kW	*HTR*	high temperature regenerator
h	enthalpy, kJ/kg	*I*	first law or energetic or thermal efficiency
HRHE	heat recovery heat exchanger	*II*	second law or exergetic efficiency
j_j	non-dimensional fraction of irreversibility rate of the j^{th} component with respect to the total exergy input, %	*isen*	isentropic
$\dot{I}_j$	irreversibility rate for j^{th} component in the thermodynamic cycle, kW	*I*	state (1 - 6) in RBC and (1 - 8) in RRCBC.
kW	kilo Watt	*in*	inlet
MPa	mega pascal		
$\dot{m}$	mass flow rate, kg/s	*LTR*	low temperature regenerator
P	absolute pressure, MPa	*MC*	main compressor
P_i	inlet stream pressure	*net*	Net (work done)
PPTD	pinch point temperature difference (in heat exchangers or regenerators), ^oC or K	*out*	outlet
$\dot{Q}$	heat rate or heat transfer per unit time basis, kW	*opt*	optimum
r_p	pressure ratio of the thermodynamic cycle or the compressor ratio (P_{max}/P_{min})		
RH	relative humidity, %	*RC*	re-compressor
S	specific entropy, kJ/kg K	*RBC*	regenerative Brayton cycle
sc	super-critical	*RRCBC*	regenerative and recompression Brayton cycle
SSSF	steady-state-steady-flow	*s*	ideal or isentropic state
T	absolute temperature, K	*w*	cooling water (sea water) in pre-cooler
$\dot{W}$	work done rate, kW	*1 to 8*	nodal state points of bottoming cycle **{Figs. 1, 2, 3(a)** and **3(b)}**
W	work done, kJ/kg	*0*	ambient or dead state
WG	water gauge (pressure)	***Greek letters***	
WHRS	waste-heat-recovery-system	ε	effectiveness
x	re-compression mass fraction (in RRCBC)	η	efficiency, %
		Δ	denotes a change in any parameter
		$\delta\overline{P}$	non-dimensional pressure drop fraction, $\delta\overline{P} = \Delta P / P_i$

Int J Syst Assur Eng Manag
DOI 10.1007/s13198-017-0629-1

ORIGINAL ARTICLE

Quantitative evaluation of web metrics for automatic genre classification of web pages

Ruchika Malhotra[1] · **Anjali Sharma**[2]

Received: 5 July 2016 / Revised: 21 February 2017
© The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2017

Abstract An additional dimension that facilitate a swift and relevant response from a web search engine is to introduce a genre class for each web page. The web genre classification distinguishes between pages by means of their features such as functionality, style, presentation layout, form and meta-content rather than on content. In this work, nineteen web metrics are identified according to the lexical, structural and functionality attributes of the web page rather than topic. The study is carried out to determine which of these attributes (lexical, structural and functionality) or its combinations, are significant for the development of web genre classification model. Also, we investigate the best web genre prediction model using parametric (Logistic Regression), non-parametric (Decision Tree) and ensemble (Bagging, Boosting) machine learning algorithms. We built forty-two genre classification models to classify web pages into Movie, TV or Music genre using a sample space data extracted from the Pixel Awards nominated and award winning websites. Our results obtained from the area under the curve analysis of these forty-two models show that the ensemble algorithms provide better performance. The rest of the models have acceptable performance, only in cases for which the lexical and structural attributes were fed in combination. Functionality metrics were found to considerably degrade the performance measure, irrespective of the algorithm used. The overall results of the study indicate the predictive capability of machine learning models for web genre classification, provided an appropriate choice is made on the selection of the input metrics.

Keywords Web metrics · Web genre classification · Machine learning · Entertainment websites

1 Introduction

As the amount of information available on the web is increasing exponentially, it is becoming more difficult to extract the desired information, closer to the user objectivity. In general, in the existing mainframe, the current search engines return the ranked list of documents depending upon textual similarity with the user query and the independent measure of each web page's importance (Choi 2001). As a result, the search still remains largely subjective, requiring further refinements to be done by the user. Certainly, one way to ease this problem is to utilize automatic genre classification of web pages into classes like personal home pages, corporate home pages, entertainment, sports etc., depending upon the information furnished. The web genre classification distinguishes between pages by means of their features such as content, structure, style, presentation layout, form and meta-content rather than based on topic. Therefore, web page genres add another dimension to the classification of web pages apart from the topic and, improve the search results returned to the user by providing genre class of a web page.

✉ Anjali Sharma
anjali1007@gmail.com

Ruchika Malhotra
ruchikamalhotra2004@yahoo.com

[1] Department of Computer Science, Delhi Technological University, Bawana Road, Delhi, India

[2] Department of Planning, Monitoring and Evaluation, CSIR-National Physical Laboratory, Dr K S Krishnan Marg, New Delhi, India

Published online: 27 May 2017

In simple terms, genre means style and, a genre class is a class of documents that are of similar type. The salient features for genre classification are the presentation and lexical properties of the document (Kennedy and Shepherd 2005). It may be noted that only textual properties are not distinctive for classification of web pages as it is a composite of HTML tags, links, navigational buttons, menus, advertisements, search boxes, texts, images, multimedia, *URL*, scripts and others. Also, the classification of web pages into the genre is a challenging task as the information is semi-structured, heterogeneous and dynamic. Hence, it is important to find appropriate features that describe a web page in the context of the genre. Even if accomplished, it is highly likely that categories within a taxonomic classification would be identified and, therefore, are to be arranged in their increasing specificity. The studies have highlighted the difficulty in automatic genre identification of websites with sparse texts (Zhu et al. 2016). For instance a case study of entertainment websites such as music, TV and movie is expected to have little textual content, but more of images and multimedia content.

A web genre classification is generally characterized by <Lexical, Structural, Functionality> tuple describing the meta-content and style, presentation layout and functionality information of the web page, respectively (Dong et al. 2008; Meyer and Stein 2004). A set of nineteen web metrics is identified in Table 1 for classification accordingly, which is quite different from the classification based on topic, itself. The choice of the sample space is currently restricted to the data from the Pixel Awards nominated and award winning websites (https://www.pixelawards.com/). The "Lexical" attributes are represented by web metrics like <meta_tags>, <meta_keywords>, <min_keyword_length>, <max_ keyword_length>, <words_alt_images> and <descriptor_words> in a web page, while the "Structural" feature is represented by web metrics like <total_links>, <total_headings>, <image_links>, <total_images>, <tables>, <total _alt_images>, , <paragraphs> and <link_headings>. The "Functionality" of the web page is represented by metrics like <scripts>, <size> and <load_time> of the page.

Beyond, earlier studies in this field of web genre classification have focused only on one classifier, which are either based on neural networks, decision trees, support vector machines, or Bayesian learning machine learning techniques. However, the effectiveness of web genre classification varies with different classifiers and data sets. Hence, in order to draw a wider generalization, it is essential to perform studies with a variety of machine learning techniques. Therefore, in this study, we compare and assess the effectiveness of parametric (Logistic Regression), non-parametric (Decision Tree) and ensemble (Bagging, Boosting) machine learning algorithms on web genre classification of Entertainment based websites.

The main objectives of our research study are summarized below:

Table 1 Web metrics selected for study as <Lexical, Structural, Functionality> tuple describing the meta-content and style, presentation layout and functionality information of the web page, respectively

Metrics	Description	Style category	Metrics	Description	Style category
meta_tags	Total number of meta tags in the page	Lexical	total_images	Total number of images on a page	Structural
meta_keywords	Total number of meta keywords in the page	Lexical	alt_images	Total number of alternate images on a page	Structural
min_keyword_length	Minimum Length of keywords	Lexical	div	Number of divisions on a page	Structural
max_keyword_length	Maximum length of keywords	Lexical	paragraphs	Total number of paragraphs on a page	Structural
descriptor_words	Total number of descriptor words on a page	Lexical	tables	Number of tables used on the web page	Structural
words_ alt_ images	Total number of words used to describe alternate images on a page	Lexical	total_headings	Total number of headings on a page	Structural
avgNoOfWordsinLink	Average Number of words in links	Structural	link_headings	Total number of links that are also headings on a page	Structural
total_links	Total number of links on a page	Structural	scripts	Total number of scripts on a page	Functionality
image_links	Total number of image links on a page	Structural	size(KBs)	Number of kilo bytes i.e. size of the page and images	Functionality
			load_time	The total time required to upload a page (in seconds)	Functionality

- We examine the effect of the <Lexical, Structural, Functionality> attributes, individually and, in combinations to classify web genre hierarchy for entertainment homepages into Music, TV and Movie. We note that previous studies have been largely based on textual data.
- We also make an extensive comparison of parametric (Logistic Regression), non-parametric (Decision Tree) and ensemble (Bagging, Boosting) machine learning algorithms on web genre classification of movie, TV and music websites.

To accomplish our objectives, the study is carried out by building forty-two genre classification models to classify web pages into Movie, TV or Music genre using a sample space data extracted from the Pixel Awards nominated and award winning websites. Systematically, the statistical analysis and pre-processing of data is first performed, and then the data is fed as input to three classification approaches, based on parametric, non-parametric and ensemble based machine learning algorithms. The forty-two genre classification models are developed and validated using 10-cross validation on the dataset. The rest of the paper is organized as follows. Section 2 describes the related work. Section 3 focuses on the data pre-processing, the evaluation metrics and model building. Section 4 describes the analysis of the results. In Sect. 5 we draw our conclusion.

2 Related work

The automatic genre classification of web pages depends upon the selection of features, automatic identification technique and the web genre palette. The feature set used in automatic genre identification represent attributes, using which genres are characterized for content, structural and functionality (Lim et al. 2005). The research in this area utilizes various machine learning algorithms for the genre prediction using feature set and existing web genre palette (Levering et al. 2008).

For instance, Bretan et al. (1998) analyzed the case data of 102 interviewers by utilizing parts of speech features along with simple text statistics like the average word length etc. The features were processed using C4.5 algorithm to mark a page into private, public/commercial, journalistic, report, interactive, discussions, link collections, FAQs and other listings. Pierre (2001) proposed super page-based classification approach considering meta tags and highlighted issues of content dependent classification. Lee and Myaeng (2002), classified a set of 765 documents into seven genre types of FAQ, Home Page, Reportage, Editorial, Research Article, Reviews and Product specification by utilizing a feature set that is based on a list of about hundred document terms in each genre class. A study based on the response of 286 interviewers and analysis of 800 documents was analyzed by Rehm (2002) by utilizing the HTML meta data, presentation tags and linguistic features for classification of web pages into academic home pages with three granularity levels. Amitay et al. (2003) utilized sixteen structural properties based on the distribution of pages and link structure for classification of websites according to its purpose. On the other hand, Eissen and Stein (2004) considered the textual information like word-frequency class, parts of speech, genre specified core vocabulary, text statistics and HTML tags for classification using SVM into eight genres. Ester et al. (2002), Kriegel and Schubert 2004) used semantic structure and local context information for the classification using Markov tree classification.

Kennedy and Shepherd (2005), used content, form and functionality tuple for classification into home pages and non-home pages with artificial neural network as a classifier. The feature set comprised of the content (common words, presence of phone number and number of meta tags), the form (CSS and HTML presentation tags like number of images and words in a page) and the functionality (script, HTML form and link tags) tuples. Their analysis found that a significant improvement could be sought in the classification, when functionality attribute was included.

Boese and Howe (2005) analyzed 342 documents on style (e.g. readability scales, parts of speech statistics), form (e.g. text statistics, HTML tags) and content (bag of words (BOW), URL, number types, closed-world sets and punctuation) features for classification into nine academic genres which were mostly text based. The study inferred that classification of genres with little textual content and with more links and images yields unsatisfactory results.

Lim et al. (2005) studied 1224 documents and classified web pages on the basis of parts of speech, URL, HTML tags, token information, most frequent function words, most frequent punctuation marks and syntactic chunks into eight non-textual genres like Personal/Public Homepages, Image Collections and eight textual genres like Research, FAQ, etc. Finn and Kushmerick (2006) investigated different textual feature sets like BOW, parts of speech frequencies, and text statistics for genre classification using C4.5 algorithm. It was concluded that a classifier can perform well with the ensemble and active learning.

Santini (2007) utilized three different feature sets, including most frequent English words, HTML tags, part of speeches, punctuation symbols, genre–specific core vocabulary and combination of layout and functionality tags for automatic genre identification. These feature sets were fed as input to SVM and Naive Bayes classifiers on

zero-to-multi genre assignment (Santini 2007). On the other hand, Kanaris and Stamatatos (2007) used variable-length character n-grams and HTML tags as feature sets for genre classification. The study validated its results on two benchmark corpora and concluded that both stylistic properties of textual data and structural information improved the SVM classifier results.

Dong et al. (2008) examined the effect of content (text visible to the user via the web browser), form (head, font, bullet, div, style, table, tr, td) and functionality (applet, script, jsp, link, form, select, option, text area, input attributes) attributes on four different web genres namely. The Information Gain Measure was used for feature selection with Naive Bayes classifier. The study inferred that combinations of content, form and functionality attributes produce better results. The results are motivating, and therefore form one of our objectives to use various combinations of lexical, form and functionality attributes in our present case study.

Chen and Choi (2008) analyzed content, URLs, HTML tags and scripts for five top-level genre classification and defined categories. The results showed that additional features help in building a better genre classification model. Lindemann and Littig (2010) presented a super-genre level classification by using structure, content and their combination as the feature set with a Naive Bayes classifier. It was found that a combination of structure and content attributes significantly improved the overall accuracy of classification.

Sharoff et al. (2010) investigated the performance of POS n-grams, character n-grams and word n-grams features on six different genre annotated datasets and found that simple character n-grams performed best. The study is comprehensive but will not work when the textual information on the page is not sufficient as in our case of entertainment based genre classification. Mason et al. (2010) also emphasized on the usage of character n-grams for genre classification. Quite differently, to improve the classification accuracy, Kumari and Reddy (2012) proposed Combined Stemming Approach, which focused on extracting genre-relevant words based on word level, linguistic features and combined stemming and stop-word elimination. Pritsos and Stamatatos (2013) used content information with character 4-gram words feature representation methods. The results showed that character 4-gram seem to provide more effective and robust models using base SVM learners and an ensemble of classifiers.

Berger et al. (2015) experimented with SVM parameters and feature sets for classification into blogs, news and forums genres of blog-like websites and identified that high precision could be achieved by using a combination of trained n-grams, structural properties and quantitative properties of web pages. Zhu et al. (2016) developed a framework to make use of On-Page and structural information of the neighboring web pages for identification of web page genre. The results were validated with SVM classifier by using two well known corpora. It was found that structural properties of web page are effective in the web page genre identification. In the recent study, Jebari (2016) proposed adaptive centroid-based classifier (ACC) where similarity of webpage is calculated with all genre centroids for multi-label classification of web pages. The study used URL extracted character n-grams, title, headings and anchors as a feature set. It was found that the ACC approach is faster and outperforms several other multi-label classifiers.

Overall, the studies discussed above provide a consensus that content, structure and functionality attributes of a web page play an important role in discovering useful information required for categorization of websites into genre. On the other hand, results of Ref. (Dong et al. 2008) and Ref. (Zhu et al. 2016) highlight that the textual content information on the page is not effective for genre identification in case of sparse textual information on the page. It may be noted that our study focus on the entertainment based genre identification, where websites have relatively little textual content in comparison with the images and multimedia content. Therefore, we focus only on the quantitative web metrics on a web page described in the form of lexical, structural and functionality attributes. Also, it has been highlighted in Ref. (Dong et al. 2008) and also in Ref. (Zhu et al. 2016), that combination of attributes perform better than single attributes. Hence, in this study, we attempt to examine the effect of the lexical, structural and functionality attributes, both individually and in combination. We also compare and assess the effectiveness of Music, TV and Movie website classifiers built using parametric, non-parametric and ensemble based machine learning algorithms.

3 Research background

3.1 Independent variables

The independent variables of this study are the nineteen web metrics collected using the tool built in Python. The tool is configured to crawl any of the URL provided in the GUI and compute all the nineteen web metrics for the page described in Table 1.

3.2 Dependent variables

The dependent variables analyzed in this study are three Entertainment related genres, Movie, TV or Music classes. The web page genre classification models are built using the statistical and machine learning algorithms to classify the web page into one of the three genres.

3.3 Performance indicator

A variety of performance indicators have been used to evaluate the predictive capability of models developed using machine learning techniques. The dataset has a disproportionate ratio of Movie, TV or Music classes and is imbalanced in nature. The Receiver Operating Characteristic (ROC) is the commonly used performance measure to deal with the imbalanced property of the dataset. The ROC curve represents the correctly predicted faulty classes (sensitivity) on the y-coordinate versus the one minus the percentage of correctly predicted non-faulty classes (1-specificity) on the x-coordinate. The optimal cutoff point that maximizes both sensitivity and specificity is determined using the ROC curve. The comparative performance analysis of each machine learning technique is evaluated using ROC curves.

The Area Under Curve (AUC) is the value of the area under the ROC Curve and its value lies between zero and one. It is a combined measure of sensitivity and specificity and is used to compute the accuracy of the predicted models. The higher the value of AUC, better is the predictive capability of the model. The AUC is insensitive to the effects of a noisy and imbalanced dataset. Hence it is advantageous to use AUC for performance evaluation of the predictive models.

3.4 Validation methods

The practical understanding of the accuracy of a model can be predicted by applying it to the different datasets other than from which it is built. Each dataset is randomly divided into 10 equal subsets. Each time one of the 10 subsets is used as the test set and the other 9 subsets are used to form a training set. The process is repeated ten times and the results from all the folds are combined to produce model result. The tenfold cross validation works in such a way that the available dataset is split into 10 portions and then a predetermined number of portion used for training the model and the rest for testing. This is iterated ten times so that each data item gets to be once in a test set.

3.5 Statistical analysis of data

It is quite rudimentary to analyze the nature and distribution of the input metrics data. We first examine the statistical quantities associated with the nineteen metrics data shown in Table 2. Our statistical analysis of the mean (μ), standard deviation (σ) and the coefficient of variation ($CV = \sigma/\mu$), clearly points to the non-normal characteristics of the data set. From the dimensionless quantity CV, which facilitates comparison of various metrics on different scales, we find that the metric data distributions don't only have a large spread of data about their mean value, but also that the relative variation of the metrics are also quite different. For instance, μ and σ computed for the metric <meta_keywords> are 4.7 and 6.3, respectively, i.e., $\mu > \sigma$, while for the metric <meta_tags> these respective quantities are 8.7 and 5.9, i.e., $\mu < \sigma$.

Further, to understand the nature of the data and how it is spread over its range, we show in Fig. 1 the frequency distribution of the nineteen metrics under consideration. It is quite evident from the distribution that the metrics data are skewed and/or are multi-modal. One may note that the factors that render skewness to a given distribution are insufficient data, presence of outliers, and/or that the distribution itself may be a multi-modal, i.e., a composite of two or more distributions. In these perspectives, a comprehensive study was carried out. Assuming that the present data is adequate to represent the problem, we

Table 2 The nineteen metrics selected for the study along with the maximum, minimum, arithmetic mean (μ) and standard deviation (σ) values from the data set

Metrics	Max	Min	μ	σ	Metrics	Max	Min	μ	σ
meta_tags	29	1	8.71	5.96	alt_images	210	0	36.93	44.56
meta_keywords	22	0	4.75	6.30	div	416	5	149.22	107.61
min_keyword_length	9	0	2.12	2.68	paragraphs	303	0	26.89	47.19
max_keyword_length	30	0	8.18	9.98	tables	4	0	0.28	0.76
descriptor_words	59	0	17.22	14.33	total_headings	81	0	20.73	17.49
words_ alt_ images	723	0	90.55	122.52	link_headings	46	0	8.83	13.09
avgNoOfWordsinLink	13.51	0.94	2.43	1.76	scripts	81	0	20.73	17.49
total_links	412	11	127.06	83.52	size(KBs)	180.29	6.20	68.44	46.86
image_links	154	1	31.12	33.50	load_time	5.52	0.3	0.92	0.86
total_images	230	1	45.36	48.29					

The statistical quantities μ and σ are calculated using the standard expressions, $\mu = (1/N) \sum x_i$, and $\sigma = (1/N)(\sum (x_i - \mu)^2)^{1/2}$, where N = total data and $\times$ an individual measure

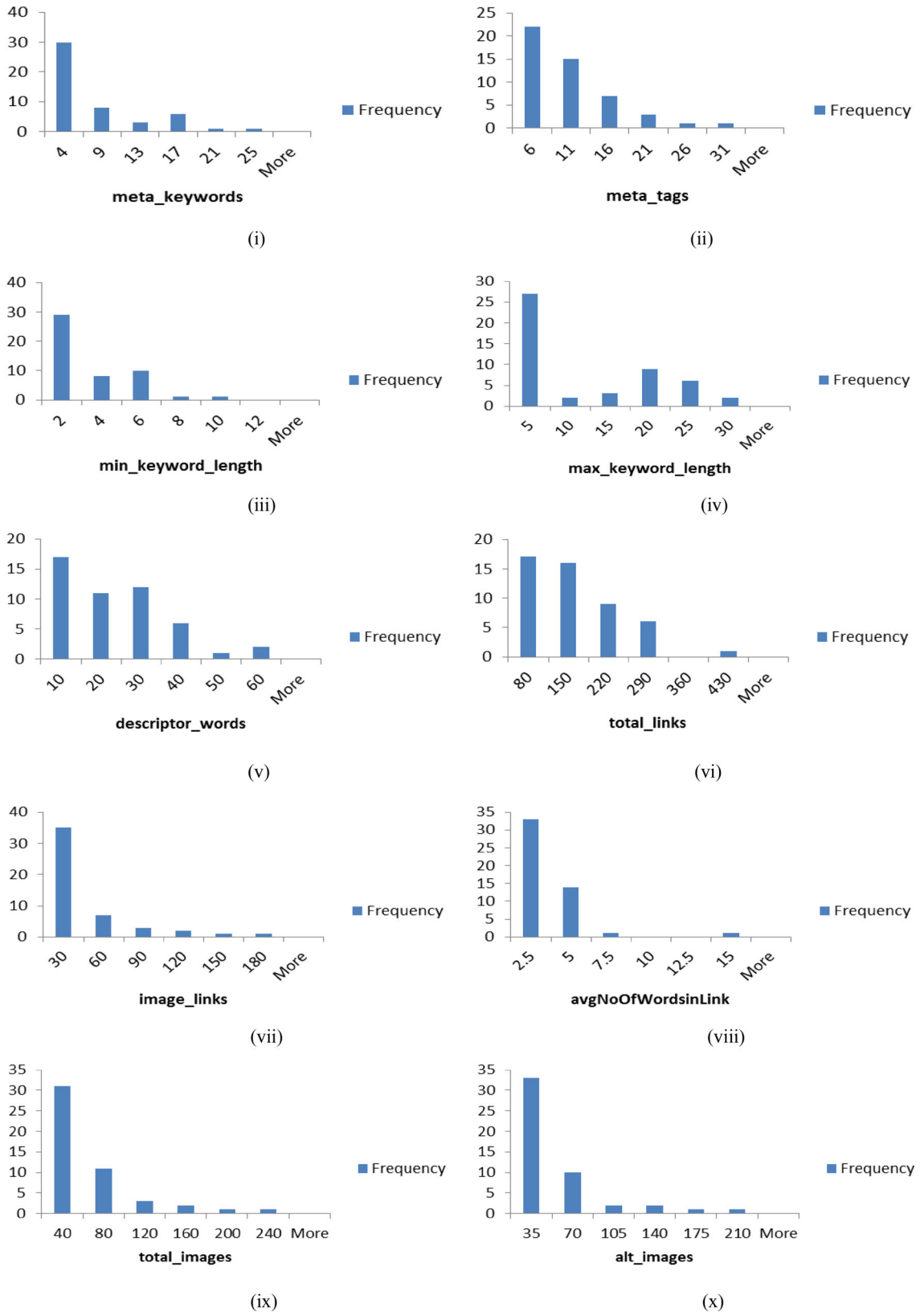


Fig. 1 Histogram plots of the metrics used for genre classification models

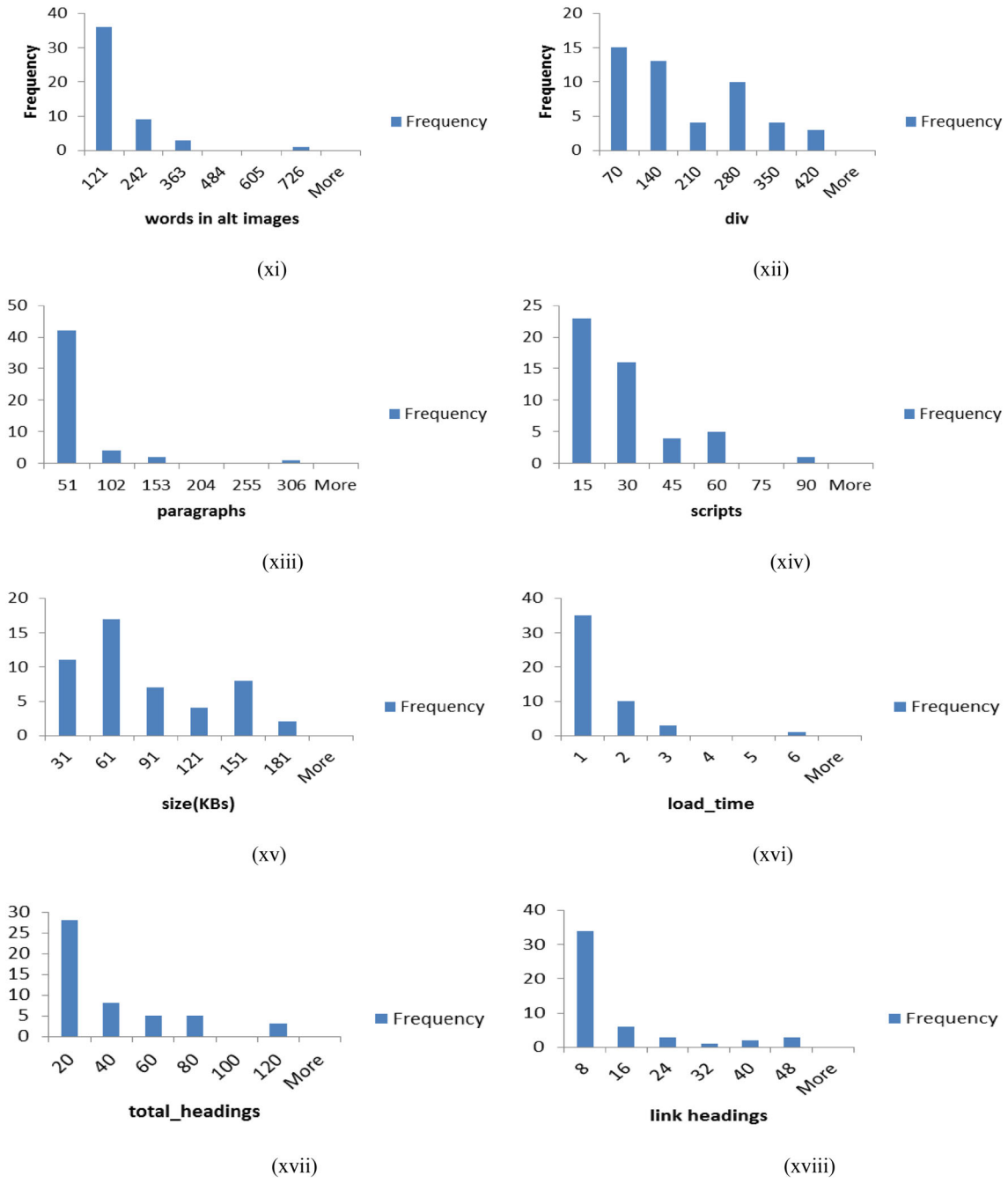


Fig. 1 continued

investigate the effects of outliers. The outliers were identified by estimating the inter-quartile range, and then appropriately removing these from the data set. By this method, we found outliers in the metrics such as <total_links>, <avgNoOfWordsinLink>, <words_alt_images>, <paragraphs>, <scripts>, <load_time> and <total_headings>.

Besides the presence of outliers, we also emphasis on whether or not the multi-modal nature of the data set, influence the final results. On a cursory glance, one could associate the non-normal characteristics of the distribution of the nature of three different categories, namely Music, Movie and TV. It is obvious that the relative importance of the metrics and its magnitude would vary in these categories, significantly. In order to substantiate our proposition on the origin of multi-modal nature in the sample data set, we show in Fig. 2, the frequency distribution of the few metrics namely, <avgNoOfWordsinLink>, <scripts> and <load_time>, refined for the three separate categories, Movie, TV and Music. It is evident from Fig. 2 that the

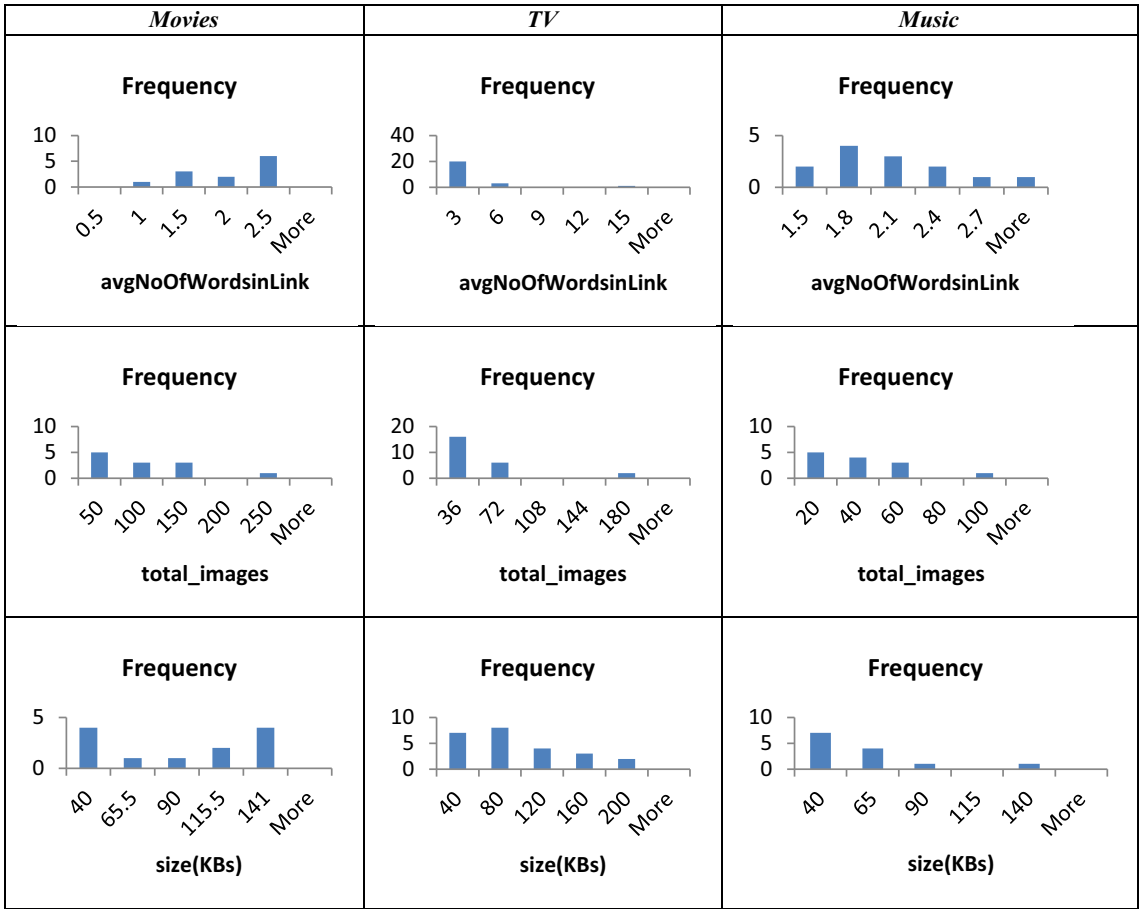


Fig. 2 Genre wise histogram plots of the metrics <avgNoOfWordsinLink>, <total_images>, <size(KB)>

distributions are very different. For instance, the distribution of metric <avgNoOfWordsinLink> in case of Movie is found to be linear, while for Music and TV it appears as skewed normal and exponential, respectively. Similarly, for the metric data <size>, it has a U-type distribution curve for the Movie category, while for Music and TV the associated distributions appear to be like exponential and skewed normal in nature. Beyond, a relevant parameter of higher importance is the range of the data. Note that the metric <total_images> ranges from 7 to 230 for Movie, while for TV and Music is different, with its range being 1–178 and 2–95, respectively. This infers that for <total_images> the Music genre is contained in the TV category, and both TV and Music, more or less contained in the Movie category.

3.6 Selection of machine learning techniques

The goal of any machine learning algorithm is to achieve low bias and low variance, resulting in good prediction performance. Model simplifications which make the mapping function easier to learn, results in bias. On the other hand, when the estimate of the mapping function changes with different training data from the population, variance emerges. Therefore, finding that both bias and variance are intimately related to each other, i.e., increasing (decreasing) bias would decrease (increase) variance, the parameterization of machine learning algorithms is often a battle to balance out bias and variance. Generally, parametric algorithms have a high bias making them fast to learn and easier to understand, but are less flexible. However, by not making assumptions they are free to learn any functional form from the training data but at the cost of time. On the other hand, non-parametric methods are often more flexible and achieve better accuracy. However, they require a larger sample space and training time. As a result, non-parametric methods seldom appear useful in complex situations, in which, introducing an assumption to simplify the complexity would lead to an overall low predictive performance, resulting in high variance as well as a high bias.

In our study, we have employed parametric (that simplify the mapping function to a known form), non-parametric (that do not make strong assumptions about the form of the mapping function) and ensemble (that create a strong classifier from a number of weak classifiers) machine learning algorithms. Although, many predictive algorithms

have been formulated in each of these machine learning categories, the scope and limitations associated with the algorithms to address a particular classification problem are intimately related to the nature of the data distribution and the underlying principles. For example, LR gives a natural probabilistic view of class distributions, but makes no assumptions about the distributions of classes in feature (attributes) space. The algorithm is also resistant to overfitting. In these perspectives, and also considering the heteroscedastic characteristics of the metric data at hand, i.e., the variability of metric variables is unequal across the range of metrics, it is anticipated that LR would yield better results to the present classification problem. Thus, a non-linear logarithmic transformation, which forms the basis of LR, becomes a choice for our investigation.

Another important type of algorithm for predictive modeling are those based on Decision Trees, which are non-parametric in nature. These algorithms suffice fast learning, hence yield faster predictions. Often, they are also quite accurate for a broad range of problems. We have used the DT algorithm, as well, for the genre prediction. In DT analysis, pre-processing of the data does not alter the final results, unlike the LR. Also, we have estimated the values using the J48 algorithm which is a C4.5 based technique.

We note that DT, in general, have high variance (low bias) and can yield more accurate predictions when used in an ensemble. In Bagging, also known as Bootstrap Aggregation, different training sets are randomly chosen from the population and are subjected to analysis under the DT classifier. The average of all such random training sets is thought to provide the best results, as such a methodology tends to compensate the increase in the error rate of any individual classifier.

Boosting is yet another ensemble technique that attempts to create a strong classifier from a set of weak classifiers. AdaBoost (or adaptive boosting), which is relatively simple and popularly used, employs short DT. In a series wise model execution, the performance of each preceding DT results is weighted. The mis-classified items are given more weight, while those which had been correctly classified are given lesser weights. Thus, models are sequentially created, each updating the weights on the training instances that affect the learning performed by the next tree in the sequence. Once all trees are built, predictions are made for the new data and the performance of each tree is weighted by its deviation from the accuracy. Another ensemble based boosting algorithm is that of logistic boost, which essentially fits a multi-nominal logistic regression model, where in each iteration one simple linear regression model is added per class.

We have utilized the open source Weka 3.8 package (Eibe Frank et al. 2016) for building the automatic web genre identification models. As tenfold cross validation was used for each model creation to get a reliable estimation of the model's capabilities in classifying the dataset. The machine learning classification techniques used in this study along with their execution settings in Weka3.8 have been described in Table 3.

4 Result and discussion

The web-genre is represented by a tuple, <lexical, structural and functionality>. The description of these tuples is given in Table 1. The first question we address is that which of these attributes or combination of attributes are significant for the web genre classification. Also, of importance is to investigate and determine the best web genre prediction model using various machine learning techniques described in Table 3. To proceed, we note that since certain web metrics have large variance and some small, we first normalize the entire data set. Since, there is no prima facie on the selection of a classifier with respect to the nature of data, we have adapted to three prediction model approaches, based on parametric (logistic regression), non-parametric (decision trees) and ensemble approach (bagging, adaptive boosting and logistic boosting). These classifiers depending on the nature, correlation and distribution of data, are found to render good predictability for certain cases, if not all. For example, the decision tree model have relatively high variance and therefore yield more accurate predictions when used in ensemble (for example, bagging), while parametric classifiers such as logistic regression makes no assumptions about data distributions and are also quite resistant to overfitting. In preparation for the classification by lexical, structural and functionality, we first sort the metric data into movies, TV and music category. This is shown in Fig. 2. We deal with data reduction, eliminating those which are lesser decisive or are noisy. Thus, for each tuple we examine the similarity of the metric data using statistical measures. We emphasize that the frequency distribution the metrics, as shown in Fig. 1, are skewed. On an empirical basis, by using the mean, median and standard deviation (see Table 2) we reduce the data set, wherever found essential, to determine the data similarity and eliminate them as input to the classifiers.

4.1 Classification by lexical

The following are the lexical attributes: <meta_keywords>, <meta_tags>, <min_keyword_length>, <max_keyword_length>, <descriptor_words> and <avgNoOfWordsinLink>. Based on their statistical measures, the data characteristics show that <min_keyword_length>,

Table 3 Description and settings of the machine learning classification techniques used in the study

Machine learning classification technique	Description and justification	Weka 3.8 settings
Logistic regression (LR)	LR gives a natural probabilistic view of class distributions, but makes no assumptions about the distributions of classes in feature space and is resistant to overfitting	Logistic -R 1.0E – 8 -M-1 -num-decimal-places 4
Decision tree (DT) J48 with C4.5	DT does not require any special pre-processing of the data and is used for a broad range of problems. The algorithm, in general, has a high variance	J48 -C 0.25 -M 2
Bagging	Bagging compensates for the increase in error rate of any individual classifier. It takes multiple samples of training set using random sampling for each individual classifier in the ensemble and the predictions are averaged to give a better estimate of the true output value	Bagging -P 100 -S 1 -num-slots 1 -I 10 –W REPTree– -M 2 -V 0.001 -N 3 -S 1 –L -1 -I 0.0
Random forest (RF)	RF is a powerful bagging algorithm using a collection of decision trees. In this approach decision trees are created so that rather than selecting optimal split points, suboptimal splits are made by introducing randomness	RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
AdaBoost (AB)	AB is used with short decision trees. Models are created sequentially one after the other, each updating the weights on the training instances that affect the learning performed by the next tree in the sequence. Once the trees are built, predictions are made for new data, and the performance of each tree is weighted by how accurate it was on the training data	AdaBoostM1 '-P 100 -I 10 -S 1 -W DecisionStump –'
LogitBoost (LB)	LB applies logistic regression techniques to the AB method for minimization of weighted least-squares error of weak learners and can handle multi-class problems	LogitBoost '-P 100 -I 10 -W DecisionStump –'

<max_keyword_length> and <descriptor_words> are similar. Henceforth, only the remaining four metrics were used for the classification analysis.The results using various classifiers are shown in Table 4. The results show that the predictability, on an average is above 60% for most of the classifiers. However, the non-parametric classifiers such as DT and J48 yield low predictability, while LR and ensemble methods yield reasonably better values. The performance of classifiers for the Music genre is lower in comparison to the classification into Movie and TV genre. In order to check whether the three omitted metrics had any significant change in the final performance measure, we also carried out the evaluation including all seven metrics. Nevertheless, we found practically no significant change in the values, with changes being in the range $\pm 3\%$. This substantiates our notion that similar data metrics could be safely ignored from the prediction calculations, using simple statistical measures such as mean, median and standard deviation.

4.2 Classification by structural

In Table 5, we show the performance of classifiers obtained using the features selected from the structural attributes as input vector. Based on the statistical similarity of the dataset, the metric <table> was eliminated from the input and therefore the selected feature set for classification was <total_links>, <image_links>, <total_images>, <alt_images>, , <paragraphs>, <total_headings> and <link heading>. The results show that the overall performance measures evaluated by means of ensemble methods were better than that of the parametric and non-parametric methods. Of these, the Movies genre was reasonably well predicted in comparison to the TV and Music categories. Further, on an average, we also note that the performance evaluation based on structural attribute was relatively poorer in comparison with that of the content metrics. Hence, it appears that the features of the form factor attribute cannot be used independently for the classification of the web genres.

4.3 Classification by functionality

The features which represent the functionality attribute are <scripts>, <size(KBs)> and <load_time>. The statistical analysis shows that these attributes have different distributions for all the three genres and hence all of these three attributes were considered for predictive analysis. In Table 6, we show the results obtained with functionality attributes. It is quite evident that the performance of all classifiers is average for all the three genres. For instance, the values obtained with decision tree and J48 is ~ 0.4, showing

Table 4 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized lexical attributes listed in Table 1

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	0.669	0.760	0.662	Bagging	**0.714**	**0.756**	**0.628**
DT	0.560	0.650	0.607	RF	0.743	0.723	0.589
J48	0.593	0.672	0.666	AB	0.631	0.695	0.609
				LB	0.673	0.735	0.637

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

Table 5 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized structural attributes listed in Table 1

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	0.489	0.390	0.498	Bagging	0.682	0.523	**0.582**
DT	0.526	0.473	0.466	RF	**0.761**	0.476	0.468
J48	0.628	0.483	0.542	AB	0.646	0.536	0.490
				LB	0.727	**0.537**	0.566

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

that these classifiers are only nominally better than a random prediction. We observed that ensemble algorithms (except AB) gave good result for Movie genre, but failed to yield better values for TV and Music. It is, however, interesting to note that LR outperforms both non-parametric and ensemble methods for Movie and Music assigns its predictive values to be 0.716 and 0.656, respectively. Although for TV, all predictors yield poor value, LR still is found to be better among all with a prediction of 0.472. This brings us to the inference that the features of functionality attribute cannot be used independently for the Movie, TV and Music genre classification on the web. The poor performance yielded by the classifiers may be due to the strong interdependent nature of its metrics data.

Thus, a summary of classification results yielded by content, form and functionality shows that independent classification based on content feature yields better values in comparison with that of form and functionality.

Table 6 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized functionality attributes listed in Table 1

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	**0.716**	**0.472**	**0.656**	Bagging	0.697	0.404	0.465
DT	0.420	0.436	0.388	RF	0.703	0.386	0.388
J48	0.412	0.407	0.462	AB	0.492	0.349	0.496
				LB	0.675	0.393	0.466

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

However, a close inspection shows that certain classifiers with form and functionality attribute serves as good predictors for certain genres. For instance, the features of the form attribute improve the performance of J48 and the ensemble algorithms over content attribute, for Movie genre. The LB classifier improved the predictability of the Movie genre by $\sim 5\%$ over content when form was used as the input. Also, in bagging, the functionality attribute yields better results for movies (0.697) in comparison to those evaluated by using the form (0.673) and content (0.652). Thus, by considering the strengths and weaknesses of the classification by content, form and functionality we attempt to evaluate the performance measures using a combination of these three attributes.

4.4 Classification by lexical and structural

The results obtained from the total of 12 metrics; four from that of lexical attribute and eight from the structural attribute, are shown in Table 7. Apart from the LR, the non-parametric and ensemble based algorithms find an improved performance value to the genres, in comparison to the case when evaluation was restricted to lexical and structural attributes, individually. The LR, however, finds considerably lower values for all three genres. Interestingly, the lexical and structural attribute combination serves much significant enhancement in the Movie genre, while only nominal changes are observed for TV and Music genres.

4.5 Classification by lexical and functionality and structural and functionality

In Tables 8 and 9, we show the results of various predictive algorithms, when both lexical and functionality, and structure and functionality were used as input. In comparison to the results shown in Tables 4, 5 and 6, the performance measures of the genres seem to have deteriorated. The values are adversely affected in the case, when the structure and functionality were taken into consideration.

Table 7 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized lexical and structural attributes

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	0.581	0.602	0.500	Bagging	0.725	0.707	0.605
DT	0.684	0.692	0.542	RF	0.747	0.711	0.564
J48	0.645	0.747	**0.672**	AB	0.693	**0.759**	0.594
				LB	**0.802**	0.718	0.590

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

Table 8 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized lexical and functionality attributes

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	0.748	**0.737**	0.551	Bagging	0.678	0.732	0.598
DT	0.560	0.650	0.607	RF	0.718	0.658	0.531
J48	0.623	0.658	0.573	AB	0.631	0.695	**0.609**
				LB	**0.759**	0.662	0.429

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

Table 9 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized structural and functionality attributes

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	0.453	0.432	**0.600**	Bagging	0.636	0.456	0.553
DT	0.526	0.473	0.466	RF	**0.755**	0.512	0.475
J48	0.613	0.485	0.530	AB	0.646	**0.536**	0.490
				LB	0.750	0.438	0.530

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

4.6 Classification by lexical, structural and functionality

So far, based on individual attributes and in pairwise combinations, we see that the pairwise combination of lexical and structural attribute render acceptable performance measure. It is obvious that functionality attribute, either by itself or in combination with lexical or structural, drastically underrates the performance of the algorithms. As a final measure, we combine all three attributes and calculate the performance measure of the genres. The results are shown in Table 10. It is observed that the performance measures are poor in comparison with that shown in Tables 4, 5 and 6. However the results obtained by choosing all the three tuples are better than the case when either structure, functionality is fed independently or in combination for classification into Movie, Music and TV genre.

These results infer that functionality attribute is not to be included in the genre prediction. As discussed above, one of the reasons may be associated with the strong correlation between the three chosen functionality metrics, namely size, load time and scripts. Besides, the underrated value of music in all above estimation owes to the fact that the metrics describing the music genre appears to be a subset of either movies and/or TV genres. Hence, it is anticipated that no specific decision boundary could be found by the algorithms to characteristically distinguish the music genre from its movies and TV counterparts.

Table 10 The AUC result for movie, TV and music web page genres estimated using the parametric, non-parametric and ensemble algorithms using normalized lexical, structural and functionality attributes

MLT	Movie	TV	Music	MLT	Movie	TV	Music
LR	0.581	0.604	0.625	Bagging	0.718	0.717	0.583
DT	0.684	0.692	0.542	RF	**0.734**	0.682	0.565
J48	0.582	**0.764**	0.642	AB	0.693	0.759	0.594
				LB	0.782	0.685	**0.645**

The best AUC value, among the machine learning techniques for Movie, TV and Music categories, is highlighted in bold

5 Conclusion

The aim of the study is to determine a better web genre prediction model amongst parametric, non-parametric and ensemble machine learning algorithms. Nineteen web metrics were collected to represent the web genre classification based on the Lexical, Structural, Functionality attributes. Overall, these describe the meta-content and style, presentation layout and functionality information of the web page of the Pixel nominated and award winning web-sites into Movies, TV and Music, respectively. The data was first statistically analyzed to identify the type of data distribution, so as to identify the similarity between various attributes. Thereafter, results of forty-two web genre classification models using parametric (LR), non-parametric (DT, RF and J48) and ensemble based (Bagging, LB and AB) machine learning algorithms were analyzed in detail. Unequivocally, the results from AUC analysis showed that the ensemble algorithms have better performance. The input data, identified in tuples, comprising of lexical, structural and functionality suggest that the functionality attribute by itself or in combination with any of the other tuples (lexical, structural or both) provide poor prediction performances. Among various combinations of the tuples, the best prediction was derived when both lexical and structural attributes serve as input to the classifiers. Also, we find a relatively poor performance in the Music genre, in comparison to the Movies and TV counterparts, irrespective of the classifier and attributes. Our analysis shows that metrics related to the Music genre appears to be a subset of either Movies and/or TV genres. Hence, the classifiers find no specific decision boundary to characteristically distinguish the Music genre from its Movies and TV genre counterparts. The overall results of the study indicate the predictive capability of machine learning models for web genre classification.

References

Amitay E, Carmel D, Darlow A, Lempel R, Soffer A (2003) The connectivity sonar: detecting site functionality by structural

patterns. In: Proceedings of the 14th conference on hypertext and hypermedia, Nottingham

Berger P, Hennig P, Schoenberg M, Meinel C (2015) Blog, forum or newspaper? Web genre detection using SVMs. In: 2015 IEEE/WIC/ACM international conference IEEE web intelligence and intelligent agent technology (WI-IAT), vol. 3, pp 64–68)

Boese ES, Howe AE (2005) Effects of web document evolution on genre classification. In: Proceedings of the CIKM'05. ACM Press, New York, pp 74–81

Bretan I, Dewe J, Hallberg A, Wolkert N, Kalgren J (1998) Web-specific genre visualization. In: Proceeding of the 30th Hawaii international conference on system science

Chen G, Choi B (2008) Web page genre classification. In: Proceedings of 2008 ACM symposium on applied computing, pp 2353–2357

Choi B (2001) Making sense of search results by automatic web page classification, In: WebNet 2001. Orlando, Florida, USA, vol. 221, pp 184–186

Dong L, Watters C, Duffy J, Shepherd M (2008) An examination of genre attributes for web page classification, In: Proceeding of the 41th annual Hawaii international conference on system sciences, pp 129–138

Eissen SM, Stein B (2004) Genre classification of web pages: user study and feasibility analysis. In: Biundo S, Fruhwirth T, Palm G (eds.) KI, pp 256–269

Ester M, Kriegel HP, Schubert M (2002) Web site mining: a new way to spot competitors, customers and suppliers in the World Wide Web. In: Proceedings of the 8th international conference on knowledge discovery and data mining, Edmonton

Finn A, Kushmerick N (2006) Learning to classify documents according to genre. J Am Soc Inform Sci Technol 7:257–262

Frank E, Hall MA, Witten IH (2016). The WEKA Workbench. Online Appendix for "Data Mining: practical machine learning tools and techniques. Morgan Kaufmann, 4th edn

Jebari C (2016) Multi-label genre classification of web pages using an adaptive centroid-based classifier. J Inf Knowl Manag 15(01):1650008

Kanaris I, Stamatatos E (2007) Web page genre identification using variable-length character n-grams. In: 19th IEEE international conference on tools with artificial intelligence vol. 7, pp 3–10

Kennedy A, Shepherd M (2005) Automatic identification of home pages on the web. In: Proceeding of the 38th annual Hawaii international conference on system sciences, HICSS-38

Kriegel HP, Schubert M (2004) Classification of websites as sets of feature vectors. In: International conference on databases and applications, Innsbruck

Kumari KP, Reddy A (2012) Performance improvement of web page genre classification. Int J Comput Appl 53(10):24–27

Lee YB, Myaeng SH (2002) Text genre classification with genre-revealing and subject-revealing features. In: Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval, ACM Press, New York, pp 145–150

Levering R, Cutler M, Yu L (2008) Using visual features for fine-grained genre classification of web pages. In: Proceeding of the 41st Hawaii international conference on system sciences

Lim CS, Lee KJ, Kim GC (2005) Automatic genre detection of web documents. In: Proceeding of natural language processing, IJCNLP 2004 Springer, pp 310–319

Lindemann C, Littig L (2010) Classification of web sites at super-genre level. In: Genres on the web, pp 211–235

Mason JE, Shepherd M, Duffy J, Keselj V, Watters C (2010) An n-gram based approach to multi-labeled web page genre classification. In: 46th Hawaii international conference on system sciences, pp 1–10

Meyer S, Stein B (2004) Genre classification of web pages:User study and feasibility analysis. In: Advances in Artificial Intelligence, Lecture Notes in Artificial Intelligence, Berlin, New York, vol. 3228, pp 256–269

Pierre JM (2001) On the automated classification of web sites. In: Linköping electronic articles in computer and information science

Pritsos D, Stamatatos E (2013) Open-set classification for automated genre identification. In: Advances in information retrieval, pp 207–217

Rehm G. (2002) Towards automatic web genre identification. In: Proceedings of the 35th Hawaii international conference on system science (HICSS02)

Santini M (2007) Automatic genre identification: towards a flexible classification scheme. In: Proceeding of the BCS IRSG symposium: future directions in information access

Santini M (2007) Characterizing genres of web pages: genre hybridism and individualization. In: Proceedings of the 40th annual Hawaii international conference on system sciences

Sharoff S, Wu Z, Markert K (2010) The web library of babel: evaluating genre collections. In: LREC

Zhu J, Xie Q, Yu SI, Wong WH (2016) Exploiting link structure for web page genre identification. Data Min Knowl Disc 30:550–575

Optical Materials 72 (2017) 31–39

Contents lists available at ScienceDirect

Optical Materials

journal homepage: www.elsevier.com/locate/optmat

Spectroscopic studies of Sm^{3+} ions activated lithium lead alumino borate glasses for visible luminescent device applications

Nisha Deopa, A.S. Rao*

Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi, 110 042, India

ARTICLE INFO

Article history:
Received 4 March 2017
Received in revised form
10 April 2017
Accepted 25 April 2017

Keywords:
Borate glasses
Samarium
Judd-Ofelt theory
Photoluminescence
Inokuti-Hirayama model
Confocal images

ABSTRACT

Photoluminescence (PL) characterization of Lithium Lead Alumino Borate (LiPbAlB) glasses doped with Sm^{3+} ions at varying concentrations have been studied by using absorption, excitation, emission, time resolved and confocal image measurements. From the absorption spectra, Judd-Ofelt (J-O) intensity parameters were evaluated and in turn used to estimate various radiative parameters for the fluorescent levels of Sm^{3+} ion doped LiPbAlB glasses. The PL spectra of Sm^{3+} ions exhibit three emission bands corresponding to the transitions $^4G_{5/2} \rightarrow {}^6H_{5/2}$, $^6H_{7/2}$ and $^6H_{9/2}$, for which the emission cross-sections and branching ratios were evaluated to know the potentialities of these materials as visible luminescent devices. The decay spectral profiles measured for $^4G_{5/2} \rightarrow {}^6H_{7/2}$ transition level were used to estimate quantum efficiency of the as-prepared glasses. The non-exponential decay curves observed for higher Sm^{3+} ion concentrations were well fitted to Inokuti-Hirayama model to understand the predominant energy transfer mechanism involved in the as-prepared glasses. CIE chromaticity coordinates and correlated color temperatures (CCT) were evaluated to understand the utility of the titled glasses in cool white light generation. The confocal images captured under 405 nm CW laser excitation show intense reddish-orange color. From the evaluated radiative parameters, emission cross-sections, quantum efficiency, CIE co-ordinates, CCT temperatures and confocal images, it was observed that LiPbAlB glass with 0.5 mol% Sm^{3+} ions are more suitable for w-LEDs and reddish-orange luminescent device applications.
© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Over the past several years, trivalent Rare Earth (RE) ions doped glassy and crystalline matrices have attracted many researchers because of their wide range commercial application in the fields of lasers, lighting, displays, drug carriers, photovoltaic devices and telecommunications [1–4]. Recently, white Light Emitting Diodes (w-LEDs) are well recognized as the next generation Solid State Lighting (SSL) technology due to their superior features such as higher efficiency, longer operating lifetime and environment friendly nature when compared with other incandescent and fluorescent sources [2,5,6]. The commercial approach to prepare w-LED is a combination of YAG: Ce^{3+} phosphor with blue InGaN LED chip. However, there is a shortfall for this approach such as low Color Rendering Index (CRI) which arise due to lack of red component, halo effect and low thermal quenching temperature [2,4]. To overcome this, a combination of RGB phosphors with n-UV LED chips could be examined to enhance the CRI value. However, this approach also encounters a drawback such as the efficiency of the red phosphor ($Y_2O_2S:Eu^{3+}$) is only about one-eighth of green ($ZnS:Cu^{+},Al^{3+}$) or blue phosphor ($BaMgAl_{10}O_{17}:Eu^{2+}$) [1,4]. Another feasibility of preparing w-LED is combining n-UV excitation source with cyan and reddish-orange light emitter. It is evident from literature that, the efforts made in this direction are moderately less [4]. This really prompted us to take up the spectroscopic investigations on Sm^{3+} ions doped glasses. Further, it is well known that, glassy materials are superior over crystalline materials from the perspective of lower production cost, simpler manufacturing technique, excellent chemical and thermal stability [5,6]. Apart from their significant role in SSL technology, reddish-orange emitting glasses provide necessary information related to the lasing properties such as transition probability, branching ratio, stimulated emission cross-sections and radiative lifetimes [7–9]. Therefore, the development of reddish-orange emitting glassy matrices with n-UV excitation sources are very important for SSL as well as laser technologies.

It is well known that, among many glass formers available, boric

* Corresponding author.
E-mail address: drsrallam@gmail.com (A.S. Rao).

http://dx.doi.org/10.1016/j.optmat.2017.04.067
0925-3467/© 2017 Elsevier B.V. All rights reserved.

acid (B_2O_3) is one of the best glass former due to its unique characteristic features such as low melting point, high transparency, good RE ion solubility and high chemical stability [10,11]. These unique properties possessed by B_2O_3 makes it as a promising host matrix to prepare luminescent devices doped with RE ions with good quantum efficiency. On the other hand, B_2O_3 possess high phonon energies (1300 cm^{-1}) due to stretching vibration of network bonding oxides, which can enhance non-radiative decay and reduces strongly the quantum efficiency [12]. Such redundant phonon energies of B_2O_3 can be reduced drastically by adding chemical constituents having relatively less phonon energies such as heavy metal oxides (HMO). Among the HMOs, Lead oxide (PbO) is well known for potential application in optoelectronics due to its high refractive index, thermal stability, wide energy gap and intermediate phonon energies [13]. Upon increasing Pb^{+2} concentration in boron network, BO_3 units converts in to BO_4 and may enhance radiative emission [14]. Presence of PbO in a glass matrix can resist the penetration of radioactive radiations because of its high atomic weight and density [13,14].

Addition of aluminum tri oxide (Al_2O_3) to borate glass matrix can improve chemical as well as thermomechanical stability and modifies the emission property of RE doped glass by acting as a network modifier [15]. It is reported in literature that, Al_2O_3 in borate glass matrix can reduce crystallization effectively and improves the stability and glass structure [16]. Al_2O_3 can form AlO_4 and AlO_6 units in coordination with the oxygen atoms within the glass structure [17]. It is reported that Al^{3+} ions in the borate glass matrix reacted in the form of triclusters and the tricluster contains three tetrahedral BO_4 along with Al_2O_3 in the form of AlO_4 with one oxygen in common [17,18]. Addition of Li_2O to a borate network enhances the transition temperature (T_g) and reduces thermal expansion coefficient [19]. Moreover Li_2O in any host matrix can move within the matrix from one side to another side and supports the replacement of Li^+ ions with other ions having same valancy. All the aforementioned scientific patronages offered by B_2O_3, PbO, Al_2O_3 and Li_2O have prompted us to prepare a glassy system namely Lithium Lead Alumino Borate (LiPbAlB) glass doped with RE ions to understand their effective usage in luminescent devices.

Among the all RE ions, Sm^{3+} ions have stimulated great importance because of their applications in diversified fields such as high density optical storage, under sea communication and visible solid state lasers [20–22]. Further, Sm^{3+} ions doped glasses possess broad emission band due to $^4G_{5/2} \rightarrow {}^6H_J$ where J = 5/2, 7/2, 9/2, 11/2. Furthermore, glasses doped with Sm^{3+} ions are also used as high power reddish orange laser sources in Micro beam Radiation Therapy (MRT) for the pre-clinical radiation treatment of cancer [23]. Sm^{3+} doped matrices are having efficient energy gap around 560 nm i.e., yellow gap. This “yellow gap” is observed in the wavelength range (590–600) nm known as amber range and assembled by either III nitrite or III phosphates laser diodes [24]. The amber LEDs are having applications in day time running lights used in vehicles and traffic signals. Further, large energy difference (7250 cm^{-1}) present between $^4G_{5/2}$ and $^6F_{11/2}$ levels of Sm^{3+} ions doped glasses prevents non-radiative decay through multi-phonon relaxation process. Therefore Sm^{3+} ions doped laser materials are superior to other RE ions doped materials because of insufficiency of lower lying excited states for the studies based on concentration quenching.

In the present work, we have described the synthesis and characterization of Sm^{3+} ions doped LiPbAlB glasses by using absorption, excitation, emission, time resolved and confocal image measurements. The main purpose of the present work is to understand the feasibility of using Sm^{3+} ions doped LiPbAlB glasses in luminescent devices such as Lasers and SSL devices.

2. Experimental

2.1. Synthesis

Sm^{3+} ions doped LiPbAlB glasses were synthesized via melt quenching techniques with the chemical composition $10Li_2O - 10PbO - (10\text{-}x)\ Al_2O_3 - 70B_2O_3 - xSm_2O_3$ (where, x = 0.1, 0.5, 1.0, 1.5, 2.0 and 2.5 mol) using appropriate amount of high purity (99.99%) Sigma Aldrich chemicals such as B_2O_3, PbO, Al_2O_3, Li_2O and Sm_2O_3. For our convenience, these glasses are designated as glass A to F. About 8 g of high purity stoichiometric composition of the batch materials mixed and grinded thoroughly in an agate mortar until to get homogeneous mixture. Such mixture collected in an alumina crucible has been subsequently heated at 1100 °C for 4–5 h in an electric furnace until the homogeneous melt was obtained. The melt was then poured on a pre-heated brass plate and then pressed quickly with another brass plate to obtain glass sample with uniform thickness. To eliminate thermal strains and air bubbles in glasses, annealing was carried out for 4 h in another electric furnace. Finally, Sm^{3+} ions doped LiPbAlB glasses were prepared to study their luminescence behavior.

2.2. Characterization

Refractive indices of Sm^{3+} ions doped LiPbAlB glasses were measured through Brewster's angle method by using He-Ne laser (650 nm line) with ±0.01 accuracy. Archimedes method was employed to measure densities of the titled glasses taking water as an immersion liquid with an accuracy of ±0.01 gm/cm^3. At room temperature (300 K) UV–vis–NIR absorption spectra were recorded by using Shimadzu 3600 spectrophotometer in 350–1450 nm wavelength range with a spectral resolution of 0.1 nm. The emission and excitation spectra for the titled glasses were recorded by using Shimadzu RF-5301-PC spectrofluorophotometer fitted with Xenon flash lamp. The time resolved PL spectra were recorded by using Edinburgh FLSP900 having xenon lamp as an excitation source with a spectral resolution of 0.1 nm. A laser scanning confocal microscope Olympus-BX51 equipped with XY piezo state was used to record confocal images of the as-prepared glasses under 405 nm CW laser.

3. Results and discussion

3.1. Absorption spectral analysis

The optical absorption spectrum of 0.5 mol % Sm^{3+} ions doped LiPbAlB glass i.e., glass B in the wavelength range from 350 to 1450 nm is depicted in Fig. 1. The absorption spectra pertaining to the remaining Sm^{3+} ions doped LiPbAlB glasses (glass A and C- F) are quite similar to glass B in band position except variation in intensity. From Fig. 1 it is observed that absorption spectrum consist of 13 bands at 362, 374, 390, 402, 416, 438, 463, 477, 946, 1079, 1227, 1372 and 1476 corresponding to the transition $^6H_{5/2} \rightarrow {}^4D_{3/2}+({}^4D,{}^6P)_{5/2}$, $^6P_{7/2}$, $^4G_{11/2}$, $^6P_{3/2}$, $({}^6P, {}^4P)_{5/2}$, $^4I_{15/2}$, $^4I_{13/2}$, $^4I_{11/2}$, $^6F_{11/2}$, $^6F_{9/2}$, $^6F_{7/2}$, $^6F_{5/2}$ and $^6F_{3/2}$ respectively. Some of the absorption bands in the UV–vis region could not appear due to the strong absorption of host glass matrix in that region. The identification and assignment of bands were done as per the data given by Carnall et al. [25]. In the present work, among all the Sm^{3+}absorption transitions observed, those which follows the selection rule $|\Delta J| \leq 6$ comes under electric dipole induced in nature and the remaining which obey the selection rule $|\Delta J| = 0, \pm 1$ comes under magnetic dipole in nature [26]. The absorption bands in the UV–vis region were found to be weak in intensity may be because of the spin forbidden selection rule as well as the overlap of different

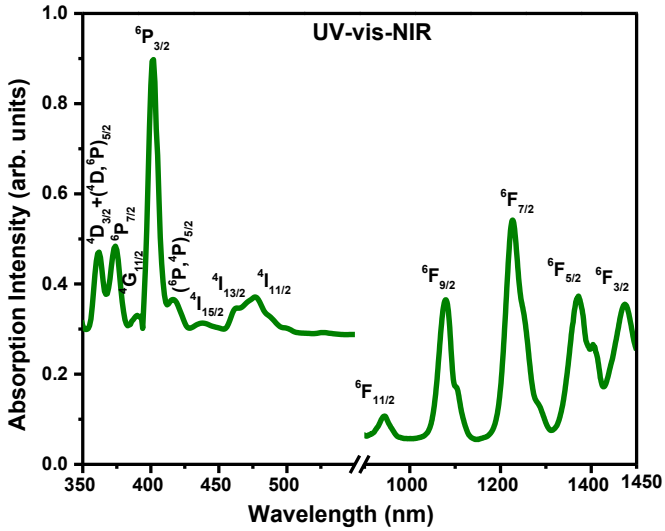


Fig. 1. Absorption Spectrum of 0.5 mol% of Sm^{3+} ions in LiPbAlB glass.

$^{2S+1}L_J$ levels in that region. The band positions observed in the NIR region are found to be intense and sharp due to the shielding effect of 5s and 5p shells to the 4f electrons. The transitions from ground state to excited states are spin allowed with $|\Delta S| = 0$ pertaining to the transitions $^6H_{5/2} \rightarrow {}^6F_J$ manifolds with J = 3/2, 5/2, 7/2, 9/2 and 11/2.

3.1.1. Nephelauxetic effect and bonding parameters

Generally, nephelauxetic effect arises due to partially filled f-shell, which is used to investigate the covalency of RE-O bond in the host matrix. Whenever the host glass is doped with RE ions, the 4f electronic orbitals of RE ions get deformed due to the influence of nephelauxetic effect. Energy level structure of RE ions constrict due to the overlap between oxygen orbital and 4f orbital, which may further induce wavelength shift. To understand the nature of bonding between Sm^{3+} ions and oxygen ligands present in the LiPbAlB glasses, nephelauxetic ratio (β) and bonding parameters (δ) were evaluated. The value of "β" can be calculated by using the following expression [27].

$$\beta = \frac{\vartheta_c}{\vartheta_a} \tag{1}$$

here, ϑ_c is the wavenumber of a particular transition of RE ion under investigation, ϑ_a is the wavenumber of the same transition of an aqua ion. In the present work, the evaluated values of "β" from glass A to F are 1.016, 1.013, 1.010, 1.008, 1.007 and 1.005 respectively. The value of "δ" can be evaluated by the following relation [28,29].

$$\delta = \frac{1 - \overline{\beta}}{\overline{\beta}} \times 100 \tag{2}$$

here, $\overline{\beta}$ is the average value of β. The "δ" parameter strongly depends upon the ligand field environment around the RE ions and can be negative or positive indicating ionic or covalent type bonding respectively. The evaluated values of "δ" from glass A to F are −1.61, −1.30, −1.04, −0.88, −0.72 and −0.55 respectively. The negative values obtained indicate the nature of bonding between Sm-O is ionic in nature, which gradually decreases as the concentration of Sm^{3+} ions increases in the titled glasses.

3.1.2. Oscillator strengths and hypersensitive transition

The energy of an absorption band of RE ion in an amorphous matrix, in other words called as oscillator strength can be evaluated by taking area under the absorption curve. Such experimental oscillator strengths (f_{exp}) can be measured by taking the areas under the absorption peaks using the following formula [30,31].

$$f_{exp} = 4.318 \times 10^{-9} \int \varepsilon(\vartheta) d\vartheta \tag{3}$$

here, $\varepsilon(\vartheta)$ is the molar extinction coefficient at a wavenumber ϑ cm^{-1}. Applying the Judd-Ofelt (J-O) [30,32] theory to all the absorption transitions observed in a host matrix, the calculated oscillator strength (f_{cal}) can be evaluated by using least square fit analysis. The necessary expression for calculating f_{cal} values was taken from the literature [33–35]. The quality of fit between f_{exp} and f_{cal} is determined by root mean square (δ_{rms}) deviation using the following relation

$$\delta_{rms} = \left[\frac{\sum \left(f_{exp} - f_{cal}\right)^2}{p}\right]^{\frac{1}{2}} \tag{4}$$

here, p represents the total number of energy levels present in the fitting procedure. The f_{exp} and f_{cal} values along with δ_{rms} are tabulated in Table 1. It can be seen from Table 1 that, 0.5 mol % Sm^{3+} ions doped LiPbAlB glass possess highest f_{exp} and f_{cal} values than the other glasses. From Table 1, it can be seen that, relatively small variation in the Sm^{3+} ions concentration can transform the oscillator strength (f_{exp}, f_{cal}) values very much and modifies the local environment around Sm^{3+} ions.

Jorgensen and Judd [30,36] reported that, the position and intensity of some of the absorption bands of RE ions are found to be very sensitive to the environment. The electric dipole transition corresponding to these bands are called hypersensitive transitions. The hypersensitive transitions follow the selection rules of quadrupole transitions $|\Delta S| = 0$, $|\Delta L| \leq 2$, $|\Delta J| \leq 2$ [36,37]. According to Judd [38], the hypersensitive transitions are mainly connected with higher values of reduced matrix element$\|U^2\|$ irrespective of $\|U^4\|$and $\|U^6\|$. Hence, the Ω_2 J-O intensity parameter plays a significant role in deciding the radiative properties. From Fig. 1 and Table 1, it is observed that the transitions corresponding to $^6H_{5/2} \rightarrow {}^6P_{3/2}$ and $^6H_{5/2} \rightarrow {}^6F_{7/2}$ are hypersensitive in nature and possess highest intensity for all the titled glasses. The degree of hypersensitivity possessed by Sm^{3+} ion can be better understood from the relative variation of Ω_2 J-O parameter for a RE ion in different amorphous matrix. Fig. 2 shows the variation of the oscillator strength of hypersensitive transition ($^6H_{5/2} \rightarrow {}^6F_{7/2}$) with Sm^{3+} ions concentration in LiPbAlB glasses.

3.1.3. Evaluation of J-O intensity parameters

The three phenomenological spectroscopic intensity parameters Ω_2, Ω_4, and Ω_6 of RE ions were obtained by applying most perceived J-O theory analysis. These parameters are very important as they can give information pertaining to the efficiency along with performance of the luminescent material. Variation of $\sum\Omega_\lambda$ with Sm^{3+} ions concentration in LiPbAlB glasses is given in Fig. 2. From Fig. 2, it is conspicuous that, $\sum\Omega_\lambda$ varies in proportionate to the intensity of hypersensitive transition. Hence their variations are following the same trend with increase in Sm^{3+} ion concentration. The Ω_2 J-O intensity parameter is related with the covalency, structural change and symmetry of the Sm-O (metal-ligand) bond. The amount of covalency is directly connected to the intensity of Ω_2 J-O parameter and nature of the hypersensitive transition [39,40]. Generally, RE

Table 1
Experimental ($f_{exp} \times 10^{-6}$), calculated ($f_{cal} \times 10^{-6}$) oscillator strengths and r.m.s deviation ($\delta_{rms} \times 10^{6}$) of Sm^{3+} ions in LiPbAlB glasses.

Transitions from $^6H_{5/2}$ →	Glass A		Glass B		Glass C		Glass D		Glass E		Glass F	
	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}	f_{exp}	f_{cal}
$^6F_{3/2}$	0.49	0.36	0.70	0.52	0.59	0.49	0.48	0.42	0.30	0.37	0.22	0.24
$^6F_{5/2}$	0.86	1.48	1.58	1.76	1.42	1.70	1.29	1.49	0.76	1.30	0.63	1.08
$^6F_{7/2}$	3.96	2.98	4.50	3.36	4.38	3.19	3.74	2.82	3.58	2.59	3.40	2.43
$^6F_{9/2}$	2.42	2.02	2.59	2.29	2.20	2.17	2.07	1.91	1.61	1.79	1.59	1.73
$^6F_{11/2}$	0.22	2.18	0.35	2.47	0.33	2.34	0.25	2.06	0.22	1.94	0.21	1.89
$^4I_{11/2}$	0.19	0.14	0.49	0.16	0.37	0.15	0.23	0.13	0.14	0.12	0.09	0.13
$^4I_{13/2}$	0.08	0.35	0.12	0.39	0.11	0.37	0.09	0.33	0.05	0.31	0.04	0.29
$^4I_{15/2}$	0.20	0.03	0.40	0.02	0.37	0.02	0.32	0.03	0.29	0.04	0.27	0.05
$(^6P, ^4P)_{5/2}$	0.25	0.48	0.49	0.58	0.45	0.56	0.41	0.49	0.38	0.43	0.29	0.36
$^6P_{3/2}$	3.78	3.42	4.24	3.82	4.06	3.69	3.87	3.26	2.98	2.83	2.09	2.39
$^4G_{11/2}$	0.08	0.02	0.15	0.19	0.11	0.19	0.06	0.02	0.04	0.03	0.02	0.05
$^6P_{7/2}$	1.58	1.26	1.89	1.39	1.70	1.32	1.65	1.17	1.54	1.09	1.43	1.06
$^4D_{3/2} + (^4D, ^6P)_{5/2}$	1.60	0.36	2.05	0.43	2.01	0.41	1.72	0.37	1.69	0.32	1.54	0.27
$\delta_{rms}(\times 10^{-6})$	±0.76		±0.82		±0.78		±0.73		±0.69		±0.67	

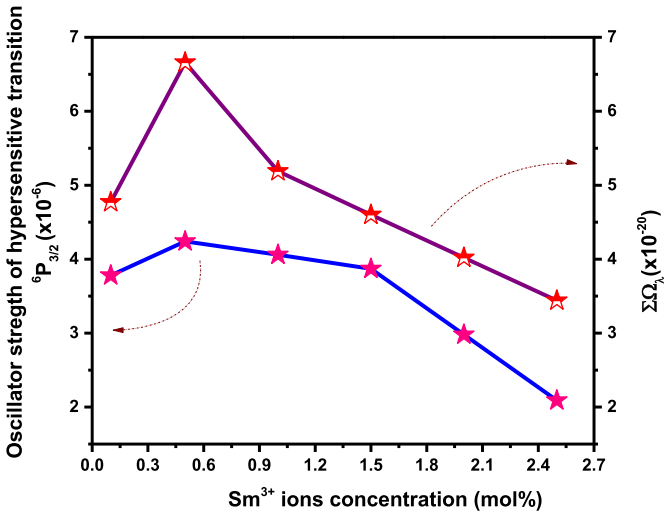


Fig. 2. Variation of intensity of hypersensitive transition and $\sum\Omega_\lambda$ with Sm^{3+} ions concentration in LiPbAlB glass.

ions in glasses are randomly disbursed in the nonequivalent sites with a wide distribution of crystal fields. In such cases, a distribution of large number of RE ions occupying sites with non-centrometric potential will contribute significant changes in Ω_2 value. When Sm^{3+} ion concentration increases in the host glass, the Ω_2 value changes suggesting the enhanced symmetry of site occupied by Sm^{3+} ions. This leads to increase in BO_4 units which further accelerate the close packing of oxygen atom around Sm^{3+} ions. Hence the variation of Sm^{3+} ion content in LiPbAlB glass can change the Ω_2 value. For the as-prepared glasses under investigation, Ω_2 is smaller than the Ω_4 and Ω_6 indicating the asymmetric behavior of ion sites which results the dominating nature of ionic bond than the covalent bond between Sm^{3+} ions and oxygen ligands. The covalent and ionic characters of the present series of glasses were estimated to be around 35.24% and 65% respectively using the equation given in our previous paper [3]. The lowest percentage value of covalent character is an evidence for the low values of Ω_2 parameter. Hence in the titled glasses, the ionic nature is more predominant than the covalent nature which is further confirmed by the higher values of Ω_4 and Ω_6 parameters. Finally this result is also in consistent with the bonding nature (δ) evaluated from the nephelauxetic ratio ($\bar{\beta}$). Further, Ω_4 and Ω_6 are associated with bulk properties such as viscosity and rigidity and are influenced by vibration transitions bound to the ligand atoms and less sensitive to the medium in which the ions are situated [34,41–43]. The J-O intensity parameters evaluated for the present glassy systems are presented in Table 2 along with some reported values [14,44–50]. It can be seen from Table 2 that J-O intensity parameters are following the same trend ($\Omega_4 > \Omega_6 > \Omega_2$) for all the titled glasses.

3.2. Photoluminescence spectral analysis

To record the photoluminescence (PL) spectra of the titled glasses, it is essential to know about the suitable excitation wavelength. Fig. 3 shows the excitation spectrum of glass B in 325–500 nm wavelength range by fixing the emission wavelength at 597 nm. The excitation spectra for the remaining glasses [glass A, C-F] are quite similar in band positions and hence were not shown in Fig. 3. The excitation spectrum of 0.5 mol% of Sm^{3+} ions doped LiPbAlB glass exhibit prominent bands at 344, 361, 374, 388, 401, 416, 439, 463 and 477 nm corresponding to the transition $^6H_{5/2} \rightarrow {}^4H_{9/2} + {}^4D_{7/2}$, $^4D_{3/2}$, $^6P_{7/2}$, $^4L_{15/2}$, $^4F_{7/2}$, $(^6P, ^4P)_{5/2}$, $^4G_{9/2}$, $^4I_{13/2}$ and $^4I_{11/2}$ respectively. In general, the phonon energy of host matrix is strongly responsible for sharpness of the observed bands while broadening in bands occurs due to Stark splitting. Among all the excitation bands, the band corresponding to $^6H_{5/2} \rightarrow {}^4F_{7/2}$ (401 nm) is very intense and is used as an excitation wavelength to record the emission spectra of Sm^{3+} ions doped LiPbAlB glasses. Fig. 4 shows the PL spectra of Sm^{3+} ions doped LiPbAlB glasses under 401 nm excitation. Three emission bands were observed at 563, 597 and 644 nm corresponding to $^4G_{5/2} \rightarrow {}^6H_{5/2}$, $^6H_{7/2}$ and $^6H_{9/2}$ transitions respectively. Among the three emission transitions, a transition

Table 2
Judd-Ofelt Parameters ($\Omega_\lambda \times 10^{-20}$cm^2) of Sm^{3+} ions in LiPbAlB glasses along with other reported values.

Glass System	Ω_2	Ω_4	Ω_6	Trend	References
Glass A	0.93	2.09	1.67	$\Omega_4 > \Omega_6 > \Omega_2$	present work
Glass B	0.86	2.45	1.52	$\Omega_4 > \Omega_6 > \Omega_2$	present work
Glass C	0.92	2.11	1.29	$\Omega_4 > \Omega_6 > \Omega_2$	present work
Glass D	0.72	1.56	1.11	$\Omega_4 > \Omega_6 > \Omega_2$	present work
Glass E	0.84	1.29	0.95	$\Omega_4 > \Omega_6 > \Omega_2$	present work
Glass F	0.70	1.00	0.86	$\Omega_4 > \Omega_6 > \Omega_2$	present work
0.5SmPbFB	0.33	3.63	2.04	$\Omega_4 > \Omega_6 > \Omega_2$	[14]
B_2O_3–Li_2O-LiF	1.15	2.28	1.34	$\Omega_4 > \Omega_6 > \Omega_2$	[44]
BMNS	4.74	7.92	5.94	$\Omega_4 > \Omega_6 > \Omega_2$	[45]
Aquo Ion	0.90	4.10	2.70	$\Omega_4 > \Omega_6 > \Omega_2$	[46]
PbO-PbF_2	1.16	2.60	1.40	$\Omega_4 > \Omega_6 > \Omega_2$	[47]
LiTFP	0.04	3.49	2.36	$\Omega_4 > \Omega_6 > \Omega_2$	[48]
TSZNB	5.21	7.85	3.70	$\Omega_4 > \Omega_6 > \Omega_2$	[49]
PbF_2–TeO_2–B_2O_3	0.21	1.42	1.87	$\Omega_6 > \Omega_4 > \Omega_2$	[50]

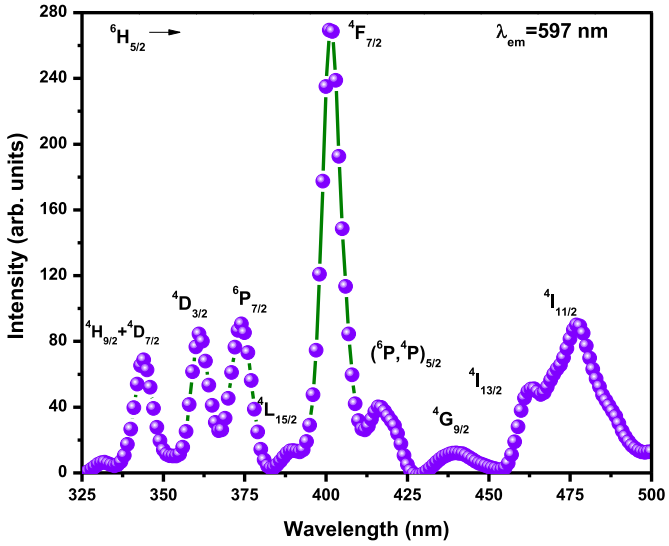


Fig. 3. Excitation spectrum of 0.5 mol% of Sm^{3+} ions in LiPbAlB glasses under 597 nm emission.

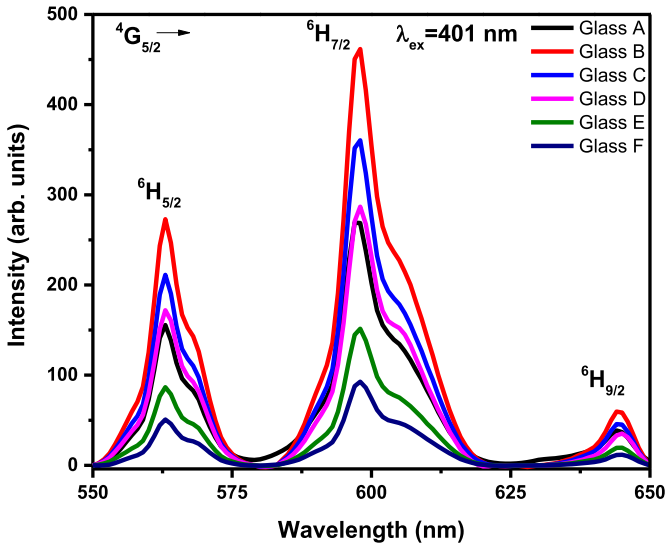


Fig. 4. PL spectra of Sm^{3+} ions in LiPbAlB glasses under 401 nm excitation.

corresponding to $^4G_{5/2} \rightarrow {}^6H_{7/2}$ (597 nm) is very intense and emits reddish orange luminescence under 401 nm excitation. From Fig. 4, it is observed that intensity of PL spectra increases with increase in Sm^{3+} ion concentration up to 0.5 mol% (glass B) and then decreases beyond showing concentration quenching. This drop in emission intensity beyond 0.5mol% may be attributed to energy transfer through cross-relaxation process between Sm-Sm ions.

The partial energy level diagram explaining emission, excitation and cross-relaxation channels for glass B can be seen from Fig. 5. As shown in Fig. 5, the energy levels situated above $^4G_{5/2}$ are closely spaced and consequently the Sm^{3+} ions excited to higher energy states above $^4G_{5/2}$ level undergoes a fast non-radiative decay to $^4G_{5/2}$ level. Further, the large energy gap existing between $^4G_{5/2}$ and lower lying $^6F_{11/2}$ level is a strong reason for the prevention of multi photon relaxation process. The possible cross-relaxation channels for Sm^{3+} ions doped LiPbAlB glasses are shown bellow

CR1: $^4G_{5/2}$, $^6H_{5/2} \rightarrow {}^6F_{11/2}$, $^6F_{5/2}$

CR2: $^4G_{5/2}$, $^6H_{5/2} \rightarrow {}^6F_{9/2}$, $^6F_{7/2}$

CR3: $^4G_{5/2}$, $^6H_{5/2} \rightarrow {}^6F_{5/2}$, $^6F_{11/2}$

CR4: $^4G_{5/2}$, $^6H_{5/2} \rightarrow {}^6F_{7/2}$, $^6F_{9/2}$

It is well know that, the cross-relaxation process occurs whenever the absorption and emission transitions matches with each other [51]. Among the aforementioned cross-relaxation channels, CR3 ($^4G_{5/2}$, $^6H_{5/2} \rightarrow {}^6F_{5/2}$, $^6F_{11/2}$) is the major energy transfer channels because $^4G_{5/2} \rightarrow {}^4F_{5/2}$ emission band possess almost same energy to that of $^6H_{5/2} \rightarrow {}^6F_{11/2}$ absorption band [52].

3.2.1. Radiative properties

From the estimated J-O parameter values, radiative parameters such as radiative transition probability (A_R), luminescence branching ratio (β_R), total transition probability (A_T) and radiative lifetimes (τ_R) were evaluated for $^4G_{5/2} \rightarrow {}^6H_{5/2}$, $^6H_{7/2}$, $^6H_{9/2}$ fluorescent levels of Sm^{3+} ions in LiPbAlB glasses using the expressions available in literature [2,17]. It is well known that a good laser host associating with larger stimulated emission cross section is very useful for achieving high gain. The stimulated emission cross section (σ_{se}) for a transition from initial to final manifold can be evaluated by using the following expression [53,54].

$$\sigma_{se}(a_j, b_{j'}) = \left[\frac{\lambda_p^4}{8\pi c n^2 \Delta\lambda_p}\right] A_R(a_j, b_{j'}) \quad (5)$$

where, λ_P represents emission peak wavelength, $\Delta\lambda_p$ is the line width (FWHM) of the transition. The values of radiative parameters A_R, β_R and A_T evaluated for Sm^{3+} doped LiPbAlB glasses are presented in Table 3. The σ_{se} along with emission peak wavelength (λ_P), effective band widths ($\Delta\lambda_P$), branching ratios (β_R & β_{exp}), gain band width (σ_{se} x $\Delta\lambda_p$) and optical gain (σ_{se} x τ_R) parameter values for the titled glasses are given in Table 4.

The branching ratio (β) is another important parameter as it describes the possibility of obtaining higher stimulated emission cross-section. For potential laser transitions, the value of "β" should be ≥ 0.5. From Table 4 it is observed that β_R & β_{exp} values are around 0.5 for $^4G_{5/2} \rightarrow {}^6H_{7/2}$ transition and are in good agreement with each other; this makes $^4G_{5/2} \rightarrow {}^6H_{7/2}$ emission channel as potential for lasing action. The higher the σ_{se} value better will be the optical gain along with low threshold. Among all the as-prepared glasses, glass B possess highest values of σ_{se} for all the three transitions $^4G_{5/2} \rightarrow {}^6H_{5/2}$, $^4G_{5/2} \rightarrow {}^6H_{7/2}$ and $^4G_{5/2} \rightarrow {}^6H_{9/2}$. The evaluated σ_{se} values for all transitions are in good agreement with some reported values [9,43,49,55–57] which are placed in Table 4.

The laser parameters such as λ_P, $\Delta\lambda_P$, β_R & β_{exp}, σ_{se} x $\Delta\lambda_p$ and σ_{se} x τ_R are very important for designing new photonic devices and for a good optical fiber these parameters should have higher values. It can be seen from Table 4 that the aforementioned parameter values are highest for glass B than the other as-prepared glasses. The reasonably higher value obtained for the aforementioned parameters for $^4G_{5/2} \rightarrow {}^6H_{7/2}$ transition of glass B makes it a potential candidate for solid state lasers as well as fiber amplifiers in the visible orange region.

3.3. Time resolved photoluminescence spectral analysis

In general, decay rates of RE doped matrices consist of radiative and non-radiative rates and radiative rate is strongly affected by the variation of local crystal symmetry at the site of RE ions, while non-radiative rates arises due to energy transfer and multi photon emission. Further, RE ions are assumed as isolated centers for lower

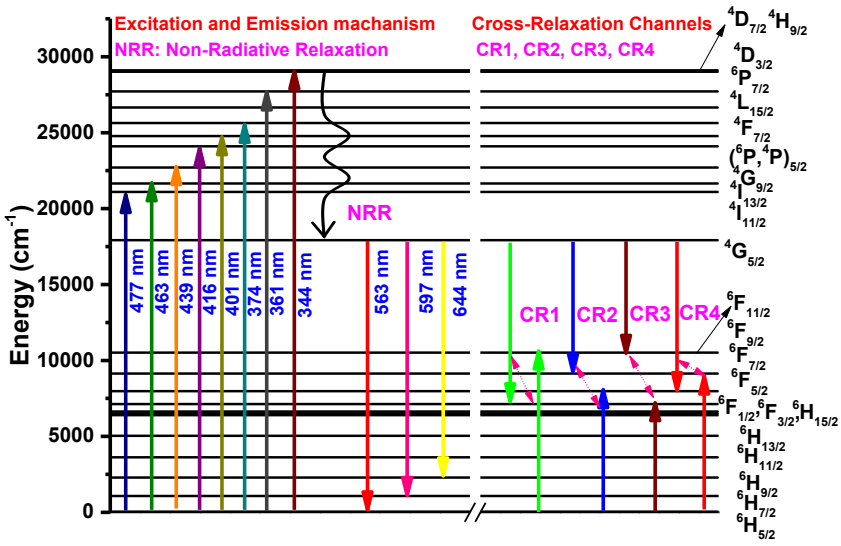


Fig. 5. Energy level diagram of 0.5 mol% Sm^{3+} ions in LiPbAlB glass with excitation, emission and cross-relaxation channels.

Table 3
Transition probability (A_R) (s^{-1}), branching ratio (β_R), total transition probability (A_T) (s^{-1}) and radiative lifetime (τ_R) (μs) for the observed emission transitions of Sm^{3+} ions in LiPbAlB glasses.

Transition	A_R	β_R	A_T	τ_R
Glass A				
$^4G_{5/2} \rightarrow {}^6H_{5/2}$	102.76	0.2494	507.87	1969
$^4G_{5/2} \rightarrow {}^6H_{7/2}$	179.74	0.4359		
$^4G_{5/2} \rightarrow {}^6H_{9/2}$	40.59	0.0976		
Glass B				
$^4G_{5/2} \rightarrow {}^6H_{5/2}$	152.92	0.2533	589.28	1697
$^4G_{5/2} \rightarrow {}^6H_{7/2}$	264.01	0.4936		
$^4G_{5/2} \rightarrow {}^6H_{9/2}$	57.98	0.0927		
Glass C				
$^4G_{5/2} \rightarrow {}^6H_{5/2}$	145.07	0.2627	625.00	1600
$^4G_{5/2} \rightarrow {}^6H_{7/2}$	251.86	0.4520		
$^4G_{5/2} \rightarrow {}^6H_{9/2}$	59.54	0.0968		
Glass D				
$^4G_{5/2} \rightarrow {}^6H_{5/2}$	137.63	0.2470	618.81	1616
$^4G_{5/2} \rightarrow {}^6H_{7/2}$	238.77	0.4237		
$^4G_{5/2} \rightarrow {}^6H_{9/2}$	61.07	0.1104		
Glass E				
$^4G_{5/2} \rightarrow {}^6H_{5/2}$	128.39	0.2618	620.34	1612
$^4G_{5/2} \rightarrow {}^6H_{7/2}$	249.56	0.4024		
$^4G_{5/2} \rightarrow {}^6H_{9/2}$	63.64	0.1155		
Glass F				
$^4G_{5/2} \rightarrow {}^6H_{5/2}$	139.08	0.2496	613.34	1630
$^4G_{5/2} \rightarrow {}^6H_{7/2}$	247.34	0.4033		
$^4G_{5/2} \rightarrow {}^6H_{9/2}$	64.29	0.1260		

Table 4
Emission peak wavelength (λ_P) (nm), effective band widths ($\Delta\lambda_P$) (nm), measured and experimental branching ratios (β_R & β_{exp}), stimulated emission cross-sections (σ_{se}) (cm^2), gain band width ($\sigma_{se} \times \Delta\lambda_P$) ($cm^3$) and optical gain parameter ($\sigma_{se} \times \tau_R$) ($cm^2$ s) for the emission transitions of Sm^{3+} ions in LiPbAlB glasses.

Spectral parameters	Glass A	Glass B	Glass C	Glass D	Glass E	Glass F
$\lambda_P = 563$ ***nm***	$^4G_{5/2} \rightarrow {}^6H_{5/2}$					
$\Delta\lambda_P$	8.12	8.43	8.34	8.22	8.07	7.98
β_R	0.25	0.26	0.24	0.23	0.21	0.19
β_{exp}	0.27	0.29	0.28	0.27	0.25	0.22
σ_{se} ($\times 10^{-22}$)	4.13	4.90	4.23	3.91	3.46	3.44
$\sigma_{se} \times \Delta\lambda_P$ ($\times 10^{-28}$)	3.36	4.13	3.53	3.21	2.78	2.70
$\sigma_{se} \times \tau_R$ ($\times 10^{-25}$)	8.14	8.32	6.77	6.31	5.56	5.64
$\lambda_P = 597$ ***nm***	$^4G_{5/2} \rightarrow {}^6H_{7/2}$					
$\Delta\lambda_P$	10.73	10.91	10.59	10.45	10.32	10.11
β_R	0.43	0.48	0.44	0.43	0.42	0.40
β_{exp}	0.48	0.54	0.47	0.44	0.41	0.38
σ_{se} ($\times 10^{-22}$)	6.80	8.11	7.23	6.64	6.50	6.03
$\sigma_{se} \times \Delta\lambda_P$ ($\times 10^{-28}$)	7.27	8.84	7.59	6.90	6.70	6.09
$\sigma_{se} \times \tau_R$ ($\times 10^{-25}$)	13.4	13.9	11.7	11.5	10.5	9.9
$\lambda_P = 644$ ***nm***	$^4G_{5/2} \rightarrow {}^6H_{9/2}$					
$\Delta\lambda_P$	5.97	6.04	6.00	5.99	5.98	5.97
β_R	0.09	0.10	0.09	0.07	0.05	0.04
β_{exp}	0.05	0.09	0.04	0.03	0.02	0.01
σ_{se} ($\times 10^{-22}$)	3.80	4.44	4.13	4.07	3.95	3.66
$\sigma_{se} \times \Delta\lambda_P$ ($\times 10^{-28}$)	2.27	2.68	2.48	2.44	2.36	2.19
$\sigma_{se} \times \tau_R$ ($\times 10^{-25}$)	7.48	7.54	6.61	6.58	6.36	5.97

concentration and decay profile from the excited state can be explained by single exponential function. Whereas for higher concentration, separation between active ion centers deceases and due to the strong interaction, excited energy transfer takes place among them. Fig. 6 shows the decay profile of Sm^{3+} doped LiPbAlB glasses showing single exponential nature for lower concentration and non-exponential for higher concentration. The non-exponential behavior arises due to the efficient energy transfer between Sm^{3+}-Sm^{3+} ions. The lifetime of Sm^{3+} doped LiPbAlB glasses decreases with increase in Sm^{3+} ions concentration and such decrease is a characteristic of concentration quenching. These results are also in consonant with concentration dependent PL spectral study. The experimental lifetime (τ_{exp}) values of $^4G_{5/2} \rightarrow {}^6H_{7/2}$ transition level for Sm^{3+} doped glasses are presented in Table 5. It is observed that experimental lifetime values are less than radiative lifetime. The discrepancy between τ_{exp} and τ_R arises due to non-radiative relaxation through cross relaxation.

The luminescence quantum efficiency (η) is another important parameter which can be calculated by using the following expression.

$$\eta = \frac{\tau_{exp}}{\tau_R} \tag{6}$$

The η(%) values are presented in Table 5 and it is observed that among all the glasses under investigation, glass B possess highest quantum efficiency.

The non-radiative decay rates (W_{NR}) can be calculated by using the following formula

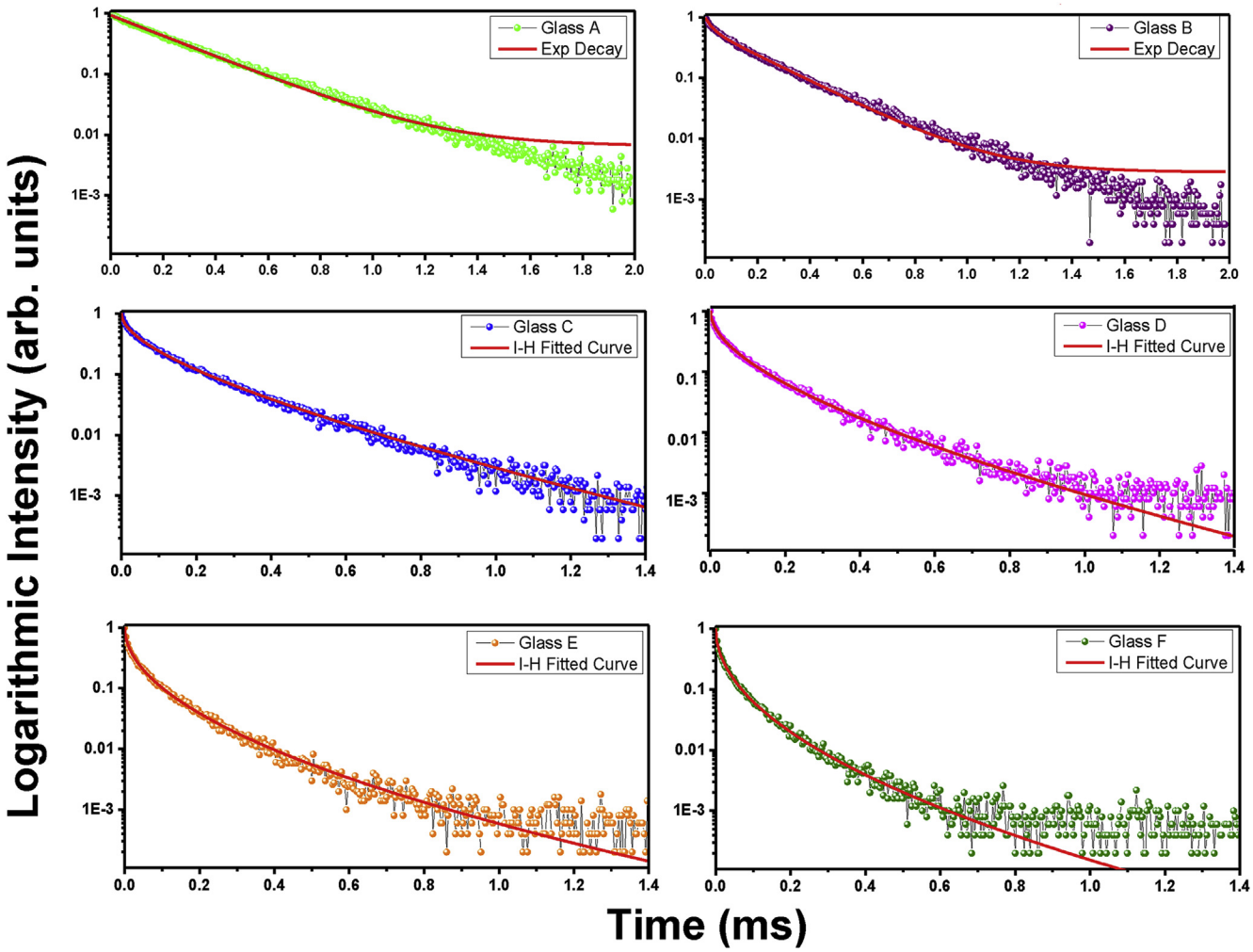


Fig. 6. The decay spectral profiles of $^4G_{5/2} \rightarrow {}^6H_{7/2}$ (597 nm) emission transition of Sm^{3+} ions in LiPbAlB glasses under 401 nm excitation along with I-H fitting.

Table 5
Experimental lifetime (τ_{exp}) (μs), energy transfer parameter (Q), critical transfer distance (R_0) (Å), donor acceptor interaction parameter (C_{DA}, 10^{-42} cm^6 s^{-1}), quantum efficiency (η) and non-radiative decay rates (W_{NR}) for $^4G_{5/2} \rightarrow {}^6H_{7/2}$ (597 nm) transition of Sm^{3+} ions in LiPbAlB glasses along with the other reported values.

Name of sample	τ_{exp}	η (%)	Q	R_0	C_{DA}	W_{NR}
Glass A	1544	78.01	–	–	–	142.36
Glass B	1492	87.91	–	–	–	79.57
Glass C	966.4	60.41	2.83	2.90	0.17	433.08
Glass D	589.3	36.46	4.59	2.98	0.21	1124.19
Glass E	386.2	23.95	14.01	3.91	1.08	2018.88
Glass F	239.8	14.71	27.08	4.54	2.62	3585.34
PbFPSm10	158	77	0.95	7.55	6.9	1427 [9]
OFBSm10	2467	83	0.24	5.15	6.7	68.65 [11]
0.5SmPbFB	1403	58	1.36	5.17	0.69	283 [14]
CFBSm0.5	2139	75	0.002	2.82	0.15	43 [21]
PPNSm05	1879	58	0.4	–	0.22	260 [41]
PKFBASm0.1	3100	72	0.005	3.17	0.31	89 [50]
LBTAF1.0	860	18	0.43	5.13	13.08	955 [59]

$$\frac{1}{\tau_{exp}} = \frac{1}{\tau_R} + W_{NR} \tag{7}$$

The W_{NR} values thus calculated are given in Table 5. From Table 5, it is observed that, the W_{NR} values are found to be least for glass B and increasing gradually up to glass F. Larger W_{NR} rates obtained at higher concentrations are also responsible for concentration quenching. The higher values of non-radiative decay rates results decrease in quantum efficiency.

3.4. Inokuti-Hirayama model analysis

The decay curves of higher Sm^{3+} ion concentrations were well fitted to Inokuti-Hirayama (I-H) model [58] in the present work taking S = 6, which indicates the interaction between Sm-Sm (donor-acceptor) ions predominantly dipole-dipole in nature. The donor atom reaching to the excited state can be de-excited either via direct energy transfer to the acceptor or through energy migration between donors until it reaches to acceptor. When the energy migration is assumed to be negligible and donor-acceptors are equally distributed in the host matrix, then the relaxation process can be understood in a better way by fitting it to the I-H model. The emission decay intensity to disclose relaxation mechanism is given by

$$I_t = I_0 \exp\left\{ -\frac{t}{\tau_0} - Q\left(\frac{t}{\tau_0}\right)^{\frac{3}{S}} \right\} \tag{8}$$

here, S can be 6 (Γ(X) = 1.77), 8 (Γ(X) = 1.43) and 10 (Γ(X) = 1.3) for dipole-dipole, dipole-quadrupole and quadrupole-quadrupole

interactions respectively. The energy transfer parameter (Q) is evaluated by the following expression

$$Q = \frac{4\pi}{3}\,\Gamma\left(1-\frac{3}{S}\right)N_0R_0^3 \tag{9}$$

here, N_0 is concentration of acceptor ions, R_0 is critical energy transfer distance and defined as the separation at which acceptor ions energy transfer rate is equal to the decay rate of donor ions. The donor-acceptor coupling constant or energy transfer probability (C_{DA}) and donor acceptor interaction parameter (W_{ET}) can be evaluated by using the following expression

$$C_{DA} = R_0^{(S)}\tau_0^{-1} \text{ and } W_{ET} = C_{DA}R_0^{(S)} \tag{10}$$

The decay profiles of higher Sm^{3+} ion concentrations in LiPbAlB glasses (glass C, D, E and F) were fitted to I-H model by taking S = 6 in this work as shown in Fig. 6. From Fig. 6 it is observed that, the values of C_{DA}, R_0 and Q increases with increase in Sm^{3+} ion concentration, confirming dipole-dipole interaction between Sm-Sm ions. This clearly demonstrates that the energy transfer via dipole-dipole interaction causes concentration quenching in the as-prepared glasses and also the same is responsible for non-exponential nature of the decay profiles. The evaluated values of C_{DA}, R_0 and Q along with some reported values are presented in Table 5 [9,11,14,21,41,50,59]. Fig. 7 shows the variation of energy transfer parameter (Q) and experimental lifetime (τ_{exp}) with Sm^{3+} ion concentration in the glasses under investigation. As Sm^{3+} ion concentration increases, the magnitude of Q increases while τ_{exp} decreases. This is quite obvious because increase in energy transfer parameter decreases the number of atoms reaching to the ground state from the excited state.

3.5. CIE chromaticity co-ordinates and confocal images

To investigate the color emitted by Sm^{3+} ions doped LiPbAlB glasses, CIE co-ordinates were evaluated from PL spectra using commission International de l'Eclaiage France CIE system [60]. The PL spectra of Sm^{3+} ions doped LiPbAlB glasses under 401 nm excitation contains three parts, the first part spread in green region while the second and third parts spread in reddish orange and red region respectively. Among these three regions, second region (reddish orange) exhibit highest intensity. Fig. 8 shows the CIE co-ordinates along with confocal images recorded for Sm^{3+} ions doped LiPbAlB glasses. It is observed from Fig. 8 that the titled glasses under 401 nm excitation falls in reddish orange region and the CIE co-ordinates

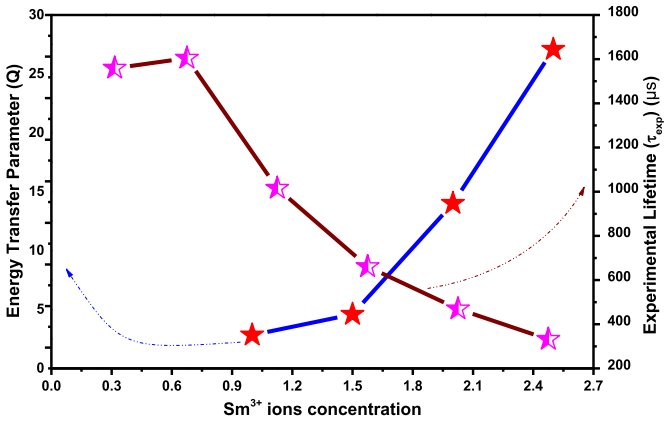


Fig. 7. Variation of energy transfer parameter (Q) and experimental lifetime (τ_{exp}) with increasing Sm^{3+} ions concentration in LiPbAlB glasses.

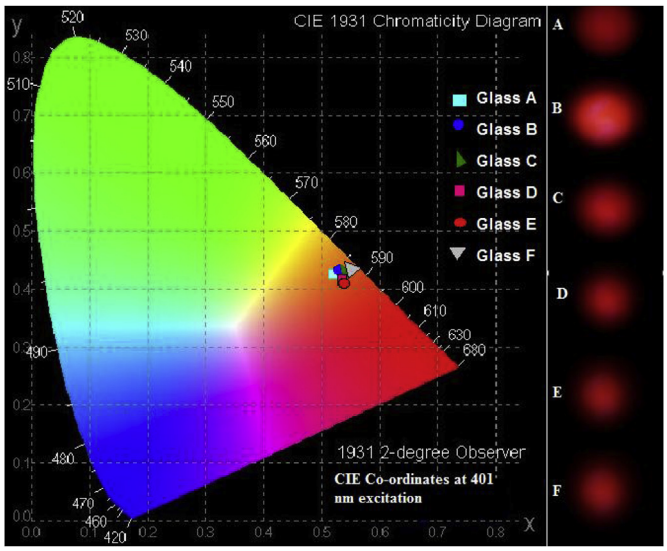


Fig. 8. CIE chromaticity coordinates and confocal images of Sm^{3+} ions in LiPbAlB glasses.

calculated for glass B are nearer to reddish orange region. This result is also in consonant with confocal images recorded under 405 nm CW laser showing relatively bright reddish orange emission for glass B. The values of CIE co-ordinates (x, y) along with CCT for Sm^{3+} ions doped LiPbAlB glasses are presented in Table 6.

4. Conclusions

LiPbAlB glasses doped with Sm^{3+} ions at different concentrations were synthesized via conventional melt quenching technique to understand their luminescence behavior. From the absorption spectral analysis it was found that the nature of bonding between Sm-O is ionic in nature, which gradually decreases with increase in Sm^{3+} ions concentration in these glasses. The PL spectra recorded for the titled glasses at 401 nm excitation show three peaks at 563, 597 and 644 nm corresponding to $^4G_{5/2} \rightarrow {}^6H_{5/2}$, $^6H_{7/2}$ and $^6H_{9/2}$ transitions respectively. Among these three transitions, $^4G_{5/2} \rightarrow {}^6H_{7/2}$ (597 nm) is having highest intensity and is suitable for reddish orange emission. The decay profiles of $^4G_{5/2} \rightarrow {}^6H_{7/2}$ transition shows single exponential for lower concentration and non-exponential for higher concentration resulting decrease in experimental lifetimes due to efficient energy transfer through cross-relaxation processes. The I-H model analysis clearly demonstrates that the predominant energy transfer process is through dipole-dipole integration. Higher values of stimulated emission cross-section, branching ratio, optical gain, gain band width and quantum efficiency observed for the transition $^4G_{5/2} \rightarrow {}^6H_{7/2}$ of glass B suggest the feasibility of using it in SSL technology and laser applications. The CIE coordinates evaluated and the confocal images captured also conclude the same thing.

Table 6
Colorimetric parameters of Sm^{3+} ions in LiPbAlB glasses under 401 nm excitation wavelength.

Glass samples	x-coordinates	y-coordinates	CCT(K)
Glass A	0.567	0.432	1877
Glass B	0.564	0.435	1901
Glass C	0.565	0.434	1896
Glass D	0.563	0.436	1916
Glass E	0.566	0.433	1889
Glass F	0.570	0.429	1846

Acknowledgement

The authors are very much thankful to Prof. G. Vijaya Prakash, IIT-Delhi and Dr. D. Haranath, CSIR-NPL, New Delhi for their experimental help in the present work.

References

[1] J. Chen, Y. Liu, L. Mei, H. Liu, M. Fang, Z. Huang, Crystal structure and temperature- dependent luminescence characteristics of $KMg_4(PO_4)_3$:Eu^{2+} phosphor for white light- emitting diodes, Sci. Rep. 5 (2015) 1–9.
[2] Ch.B. Annapurna, Sk. Mahamuda, M. Venkateswarlu, K. Swapna, A.S. Rao, G. Vijaya Prakash, Dy^{3+} ions doped single and mixed alkali fluoro tungsten tellurite glasses for LASER and white LED applications, Opt. Mater 62 (2016) 569–577.
[3] N. Deopa, A.S. Rao, Sk. Mahamuda, M. Gupta, M. Jayasimhadri, D. Haranath, G. Vijaya Prakash, Spectroscopic studies of Pr^{3+} doped lithium lead alumino borate glasses for visible raddish orange luminescent device applications, J. Alloys Compd. 708 (2017) 911–921.
[4] A.K. Vishwakarma, M. Jayasimhadri, Pure orange color emitting Sm^{3+} doped $BaNb_2O_6$ phosphor for solid-state lighting applications, J. Lumin 176 (2016) 112–117.
[5] Sk. Mahamuda, K. Swapna, P. Packiyaraj, A.S. Rao, G. Vijaya Prakash, Lasing potentialities and white light generation capabilities of Dy^{3+} doped oxy-fluoroborate glasses, J. Lumin 153 (2014) 382–392.
[6] B. Tian, B. Chen, Y. Tian, X. Lie, J. Zhang, J. Sun, H. Zhong, L. Cheng, S. Fu, H. Zhong, Y. Wang, X. Zhang, H. Xia, R. Hua, J. Mater. Chem. C 1 (2013) 2338–2344.
[7] Y. Zheng, B. Chen, H. Zhong, J. Sun, L. Cheng, X. Li, J. Zhang, Y. Tian, W. Lu, J. Wan, T. Yu, L. Huang, H. Yu, H. Lin, Optical transition, excitation state absorption and energy transfer study of Er^{3+}, Nd^{3+} single-doped, and Er^{3+}/Nd^{3+} codoped tellurite glasses for mid-infrared laser applications, J. Am. Ceram. Soc. 94 (2011) 1766–1772.
[8] A.S. Rao, J.L. Rao, R.R. Reddy, T.V.R. Rao, Electron paramagnetic resonance and optical absorption spectra of Cr (III) ions in alkali cadmium borosulphate glasses, Opt. Mater 4 (1995) 717–721.
[9] C.R. Kesavulu, C.K. Jayasankar, Spectroscopic properties of Sm^{3+} ions in lead fluorophosphate glasses, J. Lumin 132 (2012) 2802–2809.
[10] Y. Zhang, C. Lu, L. Sun, Z. Xu, Y. Ni, Influence of Sm_2O_3 on the crystallization and luminescence properties of boroaluminosilicate glasses, Mater. Res. Bull. 44 (2009) 179–183.
[11] Sk. Mahamuda, K. Swapna, M. Venkateswarlu, A.S. Rao, S. Shakya, G. Vijaya Prakash, Spectral characterisation of Sm^{3+} ions doped Oxy-fluoroborate glasses for visible orange luminescent applications, J. Lumin 154 (2014) 410–424.
[12] S.S. Sundari, K. Marimuthu, M. Sivraman, S.S. Babu, Composition dependent structural and optical properties of Sm^{3+}-doped sodium borate and sodium fluoroborate glasses, J. Lumin 130 (2010) 1313–1319.
[13] A.G.S. Filho, J.M. Filho, F.E.A. Melo, M.C.C. Custodio, R. Lebullenger, A.C. Hernandes, Optical properties of Sm^{3+} doped lead fluoroborate glasses, J. Phys. Chem. Solids 61 (2000) 1535–1542.
[14] S. Arunkumar, K. Marimuthu, Concentration effect of Sm^{3+} ions in B_2O_3 – PbO – PbF_2 – Bi_2O_3 – ZnO glasses – structural and luminescence investigations, J. Alloys Compd. 565 (2013) 104–114.
[15] J.E. Shelby, Introduction to Glass Science and Technology, The Royal Society of Chemistry, Cambridge, 1997.
[16] L.V. Wullen, W.M. Warmuth, ^{11}B MAS NMR spectroscopy for characterizing the structure of glasses, Solid State Nucl. Magn. Reson 2 (1993) 279–284.
[17] K. Swapna, Sk. Mahamuda, A.S. Rao, S. Shakya, T. Sasikala, D. Haranath, G. Vijaya Prakash, Optical studies of Sm^{3+} ions doped zinc alumino bismuth borate glasses, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 125 (2014) 53–60.
[18] H. Scholze, Glass: Nature, Structure and Properties, Springer, New York, 1991.
[19] M.C. Eckersley, P.H. Gaskell, A.C. Barnes, P. Chieux, Structural ordering in a calcium silicate glass, Nature 335 (1988) 525–527.
[20] Y. Wang, C. Lin, H. Zheng, D. Sun, L. Li, B. Chen, Fluorescent and chromatic properties of visible -emitting phosphor KLa(MoO4)2:Sm^{3+}, J. Alloys Compd. 559 (2013) 123–128.
[21] J.S. Kumar, K. Pavani, T. Sasikala, A.S. Rao, N.K. Giri, S.B. Rai, L.R. Moorthy, Photoluminescence and energy transfer properties of Sm^{3+} doped CFB glasses, Solid State Sci. 13 (2011) 1548–1553.
[22] C.K. Jayasankar, V. Venkatramu, P. Babu, T. Tröster, W. Sievers, G. Wortmann, W.B. Holzapfel, High-pressure fluorescence study of Sm^{3+}-doped borate and fluoroborate glasses, J. Appl. Phys. 97 (2005) 93523–93527.
[23] D.N. Slatkin, P. Spanne, F.A. Dilmanian, M. Sandborg, Microbeam radiation therapy, Med. Phys. 19 (1992), 1395-1234.
[24] M.R. Krames, O.B. Shchekin, R. Mueller-Mach, G.O. Mueller, L. Zhou, G. Harbers, M.G. Craford, Status and future of high-power light-emitting diodes for solid-state lighting, J. Disp. Technol. 3 (2007) 160–175.
[25] W.T. Carnall, P.R. Fields, K. Rajnak, Electronic energy levels in the trivalent lanthanide aquo ions. I. Pr^{3+}, Nd^{3+}, Pm^{3+}, Sm^{3+}, Dy^{3+}, Ho^{3+}, Er^{3+}, and Tm^{3+}, J. Chem. Phys. 49 (1968) 4424–4442.
[26] A. Thulasiramudu, S. Buddhudu, Optical characterization of Sm^{3+} and Dy^{3+}: ZnO – PbO – B_2O_3 glasses, Spectrochim. Acta - Part A 67 (2007) 802–807.
[27] C.K. Jorgensen, The nephelauxetic series, Prog. Inorg. Chem. 4 (1962) 73.
[28] W.T. Carnall, K.A. Gschneidner, L. Eyring (Eds.), Hand Book of Physics and Chemistry of Rare-earths, 1987.
[29] S.P. Sinha, Complexes of the Rare Earths, Pergamon Press, Oxford, 1966.
[30] B.R. Judd, Optical absorption intensities of rare-earth ions, Phys. Rev. 127 (1962) 750–761.
[31] G. Lakshminarayana, S. Buddhudu, Spectral analysis of Eu^{3+} and Tb^{3+}: B_2O_3 – ZnO – PbO glasses, Mater. Chem. Phys. 102 (2007) 181–186.
[32] G.S. Ofelt, Intensities of crystal spectra of rare earth ions, J. Chem. Phys. 37 (1962) 511–520.
[33] P. Babu, C.K. Jayasankar, Spectroscopic properties of Dy^{3+} ions in lithium borate and lithium fluoroborate glasses, Opt. Mater 15 (2000) 65–79.
[34] R. Praveena, V. Venkatramu, P. Babu, C.K. Jayasankar, Fluorescence spectroscopy of Sm $^{3+}$ ions in P_2O_5 – PbO – Nb_2O_5 glasses, Phys. B 403 (2008) 3527–3534.
[35] C.S. Rao, C.K. Jayasankar, Spectroscopic and radiative properties of Sm^{3+}-doped K – Mg – Al phosphate glasses, Opt. Commun. 286 (2013) 204–210.
[36] C.K. Jorgensen, B.R. Judd, Hypersensitive pseudoquadrupole transitions in lanthanides, Mol. Phys. 8 (1964) 281–290.
[37] S. Tanabe, T. Ohyagi, N. Soga, T. Hanada, Compositional dependence of Judd-Ofelt parameters of Er^{3+} ions in alkali-metal borate glasses, Phys. Rev. B 46 (1992) 3305–3310.
[38] S. Misra, Energy interaction parameters and intensity analysis of praseodymium and neodymium complexes, J. Sci. Ind. Res. 44 (1985) 366–375.
[39] K. Maheshvaran, K. Linganna, K. Marimuthu, Composition dependent structural and optical properties of Sm^{3+} doped boro-tellurite glasses, J. Lumin 131 (2011) 2746–2753.
[40] Y.K. Sharma, S.S.L. Surana, R.P. Dubedi, V. Joshi, Spectroscopic and radiative properties of Sm^{3+} doped zinc fluoride borophosphate glasses, Mater. Sci. Eng. B 119 (2005) 131–135.
[41] P. Srivastava, S.B. Rai, D.K. Rai, Optical properties of Sm^{3+} doped calibo glass with addition of lead oxide, Spectrochim. Acta - Part A 60 (2004) 637–642.
[42] Sk. Mahamuda, K. Swapna, P. Packiyaraj, A.S. Rao, G. Vijaya Prakash, Visible red, NIR and Mid-IR emission studies of Ho^{3+} doped zinc alumino bismuth borate glasses, Opt. Mater 36 (2013) 362–371.
[43] K.S.V. Sudhakar, M.S. Reddy, L.S. Rao, N. Veeraiah, Influence of modifier oxide on spectroscopic and thermoluminescence characteristics of Sm^{3+} ion in antimony borate glass system, J. Lumin 128 (2008) 1791–1798.
[44] K.K. Mahato, D.K. Rai, S.B. Rai, Optical studies of Sm^{3+} doped oxyfluoroborate glass, Solid State Commun. 108 (1998) 671–676.
[45] I.A. Rayappan, K. Selvaraju, K. Marimuthu, Structural and luminescence investigations on Sm^{3+} doped sodium fluoroborate glasses containing alkali/alkaline earth metal oxides, Phys. B 406 (2011) 548–555.
[46] R. Reisfield, C.K. Jorgenson, Handbook on the Physics and Chemistry of Rare Earths, vol. 9, Elsevier, Amsterdam, 1987.
[47] P. Nachimuthu, R. Jagannathan, V.N. Kumar, D.N. Rao, Absorption and emission spectral studies of Sm^{3+} and Dy^{3+} ions in PbO-PbF_2 glasses, J. Non. Cryst. Solids 217 (1997) 215–223.
[48] M. Jayasimhadri, L.R. Moorthy, S.A. Saleem, R.V.S.S.N. Ravikumar, Spectroscopic characteristics of Sm^{3+}-doped alkali fluorophosphate glasses, Spectrochim. Acta - Part A 64 (2006) 939–944.
[49] O. Ravi, C.M. Reddy, L. Manoj, B.D.P. Raju, Structural and optical studies of Sm^{3+} ions doped niobium borotellurite glasses, J. Mol. Struct. 1029 (2012) 53–59.
[50] B.C. Jamalaiah, M.V.V. Kumar, K.R. Gopal, Fluorescence properties and energy transfer mechanism of Sm^{3+} ion in lead telluroborate glasses, Opt. Mater 33 (2011) 1643–1647.
[51] T. Suhasini, J.S. Kumar, T. Sasikala, K. Jang, H. Sueb, M. Jayasimhadri, J.H. Jeong, S.S. Yi, L.R. Moorthy, Absorption and fluorescence properties of Sm^{3+} ions in fluoride containing phosphate glasses, Opt. Mater 31 (2009) 1167–1172.
[52] J.S. Kumar, K. Pavani, A.M. Babu, N. Kumar, S.B. Rai, L.R. Moorthy, Fluorescence characteristics of Dy^{3+} ions in calcium fluoroborate glasses, J. Lumin 130 (2010) 1916–1923.
[53] R. Cases, M.A. Chamarro, R. Alcala, V.D. Rodriguez, Optical properties of Nd^{3+} and Dy^{3+} ions in ZnF_2-CdF_2 based glasses, J. Lumin 48 (1991) 509–514.
[54] W.F. Krupke, Induced-emission cross sections in neodymium laser glasses, IEE J. Quantum Electron 10 (1974) 450–457.
[55] A. Kumar, D.K. Rai, S.B. Rai, Optical properties of Sm^{3+}ions doped in tellurite glass, Spectrochim. Acta - Part A 59 (2003) 917–925.
[56] C.M. Reddy, G.R. Dillip, B.D.P. Raju, Spectroscopic and photoluminescence characteristics of Dy^{3+} ions in lead containing sodium fluoroborate glasses for laser materials, J. Phys. Chem. Solids 72 (2011) 1436–1441.
[57] K. Lim, N. Vijaya, C.R. Kesavulu, C.K. Jayasankar, Structural and luminescence properties of Sm^{3+} ions in zinc fluorophosphate glasses, Opt. Mater 35 (2013) 1557–1563.
[58] M. Inokuti, F. Hirayama, Influence of energy transfer by the exchange mechanism on donor luminescence, J. Chem. Phys. 43 (1965) 1978–1988.
[59] B.C. Jamalaiah, J.S. Kumar, A.M. Babu, T. Suhasini, L.R. Moorthy, Photoluminescence properties of Sm^{3+} in LBTAF glasses, J. Lumin 129 (2009) 363–369.
[60] F.W. Billmeyer, M. Saltzman Jr., Principles of Colour Technology Wiley, 1981. New York.

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Terahertz pulses in a periodic plasma channel via modulated electron beam

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 J. Phys.: Conf. Ser. 836 012001

(http://iopscience.iop.org/1742-6596/836/1/012001)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 14.139.251.107
This content was downloaded on 06/06/2017 at 11:18

Please note that terms and conditions apply.

You may also be interested in:

Terahertz radiation generation by using modulated electron beam with plasma wave wiggler
J Panwar, S C Sharma and R Sharma

Effect of pulse slippage on beat wave THz generation in a rippled density magnetized plasma
Manoj Kumar, Satish Kumar Rajouria and Magesh Kumar K K

A tunable terahertz radiation source based on a surface wave transformed into Cherenkov radiation in a subwavelength array
Ping Zhang, Min Hu, Renbin Zhong et al.

Scattering and propagation of terahertz pulses in random soot aggregate systems
Li Hai-Ying, Wu Zhen-Sen, Bai Lu et al.

Tailoring the Spectra of Terahertz Emission from CdTe and ZnTe Electro-Optic Crystals
Minwoo Yi, Kang Hee Lee, Inhee Maeng et al.

Generation of terahertz radiation by focusing femtosecond bichromatic laser pulses in a gas or plasma
P A Chizhov, Roman V Volkov, V V Bukin et al.

Terahertz Pulse Detection Using Rotary Optical Delay Line
Geun-Ju Kim, Seok-Gy Jeon, Jung-Il Kim et al.

Generation of frequency-tunable pulsed terahertz radiation by a Cr : forsterite laser system with an acoustooptical control of the pulse temporal profile
A.V. Ovchinnikov, O.V. Chefonov, V.Ya. Molchanov et al.

IOP Conf. Series: Journal of Physics: Conf. Series **836** (2017) 012001 doi:10.1088/1742-6596/836/1/012001

Terahertz pulses in a periodic plasma channel via modulated electron beam

Pratibha Malik[a], Suresh C Sharma, Rinku Sharma

Department of Applied Physics, Delhi Technological University, Delhi-110042, India

Email: [a)]pratibhamalik27@gmail.com

Abstract.The proposed scheme offers a promising way to study the effect of modulated relativistic electron beam in a periodic plasma channel (wiggler). The fluid equations model is used to find out the numerical behaviour of the radiation wave. In this scheme, a modulated relativistic electron beam makes possible to generate and amplify the high power terahertz (THz) radiation pulses. Using this model, we have analyzed and calculated the nonlinear dispersion relation of the radiation wave. The radiation frequency of THz wave can be tuned by varying the wiggler wavelength and beam energy. The excitation of the radiation wave is being developed with the resonance between the modulated frequency of electron beam and radiation wave in the periodic plasma channel (wiggler field) via Cerenkov interaction along the axial direction of channel, which provides the maximum enhancement in the growth rate of the radiation wave. Additionally, the growth rate of the unstable mode of the radiation wave depends on beam energy and modulation index, and becomes larger as modulation index approaches unity.

1. Introduction

Terahertz far-infrared radiation (THz) is considered to be the significant cutting edge of the electromagnetic spectrum for the last few decades, with various unique utility in high resolution imaging with medical application, security inspection, and remote sensing [1-4]. The THz radiation can be generated using various schemes such as ultra-short laser pulses [5] and quantum cascade lasers [6].

The conventional electromagnetic wiggler based FEL has been verified an effective source of generating short wavelength electromagnetic radiation but due to difficulties occurring in the designing of the practical implementations of such wigglers are limited. To overcome this limitation new plasma wigglers[7- 8] based FELs are introduced.The injection of an electron beam in a periodically system is a crucial example for the development of narrative radiation sources such as free electron lasers (FEL). In this paper, we present a schemeto generate and amplify Terahertz radiation using under-denserippled plasma wiggler.

2. Model

Consider a density modulated relativistic electron beam of velocity $v_b^0\hat{z}$ and density $n_{b1}(n_b^0+\Delta n_b^0\exp[-i\omega t+i\omega z/v_b^o])$ is being introduced in ripple plasma wiggler of density $n_w(n_p^0+n_w^p\exp[i\vec{w}.\vec{r}])$ with wave vector $\vec{w}$.Here, the terms n_p^0 and n_w^p are the density of plasma electrons at equilibrium and perturbation, respectively. This non-linear interaction excites and

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

amplifies the THz electromagnetic wave $(\omega_1 = 2\gamma_0^2 w v_b^0, \vec{k}_1)$ co-propagating with beam wave($(\omega_1, \vec{k}_b)$. The response offered by beam electrons and plasma electrons on account of THz radiation and beam wave is governed by the equation of motion

$$\frac{\partial}{\partial t}(\gamma_1 \vec{v}_{b1}) + (\vec{v}_{b1}.\nabla)\gamma_1 \vec{v}_{b1} = -\frac{e\vec{E}_T^1}{m} - \frac{e\vec{E}_b^1}{m} - \frac{e}{mc}(v_{b1} \times \vec{B}_T), \tag{1}$$

$$\frac{\partial n_{b1}}{\partial t} + \nabla.(n_{b1}\vec{v}_{b1}) = 0. \tag{2}$$

After solving the equations (1) and (2) further, we get the perturbed velocity and density of the beam and plasma electrons, respectively. Consequently, this produces the nonlinear current density for THz radiation wave and beam wave. This nonlinear current density stimulate the excitation of the unstable radiation wave

$$\vec{J}_n^l = (\vec{J}_n^l)_T + (J_n^l)_{beam} \tag{3}$$

Furthermore, using the Maxwell's third and fourth equations, we solve the wave equations and get the nonlinear dispersion relation and growth rate in unstable mode of radiation wave as follows

$$\lambda_w = \frac{4\pi c\omega_1(\gamma_o^2 - 1)\cos(\psi)}{\omega_1^2 + \gamma_o^2\omega_p^2}, \tag{4}$$

$$\beta = \left[\frac{3}{16}\right]^{\frac{1}{2}}\left[\frac{\omega_b^2\omega_w^4 w_x^2}{\gamma_o^3 c^2(\omega_p^2 + k_1^2c^2)^{1/2}(k_b^2 - k_1^2)^2}\left((1+\Delta+(1+\Delta)\gamma_o^2\frac{w_x^2}{k_{bz}^2}) + \Delta\frac{k_o}{k_{bz}}\right)\right]^{\frac{1}{3}}. \tag{5}$$

where the terms ω_b , ω_w and ω_p are represented as beam plasma frequency, plasma wiggler frequency and electron plasma frequency, respectively in equation (5). This relation shows that the growth rate in the unstable mode of THz radiation wave depends on density of beam electrons or beam current, plasma electrons and inversely proportional to radiation frequency of the wave. Additionally, the above results can be extended to find out the efficiency of the radiation system.

3. Results and Discussions

The results show that the amplification in growth rate of the unstable mode of wave is the sensitive function of the plasma density, beam current and modulation index. Using equation (4),the variation of radiation frequency versus plasma wiggler period is shown in figure 1.This figure illustrates that the wiggler period initially increases and then sharply starts decreasing as frequency increases further. In addition, as the plasma density increases more, the wiggler period becomes shorter and it could be more resonant to generate and amplify the coherent THz radiation. Using equation (5), figure 2 depicts the variation of growth rate of the radiation wave with beam density. This shows that growth rate increases as beam density increases. This result is in good agreement with the results obtained by[9-10]. From figure 3, it can be seen that the growth rate of the wave increases with modulation index but decreases as beam energy increases.

Using equation (4), in figure 4, we observe the variation of wiggler period with ripple angle and plasma density. This figure illustrates that the wiggler period decreases as ripple angle and plasma density increases and becomes sharper to amply the THz radiation emission.

4. Conclusions

In summary, the interaction of pre-bunched relativistic electron beam with plasma kind of wiggler has been considered as one of the important alternative scheme to amplify the THz radiation emission. The present scheme reveals that the amplification of THz radiation wave depends on the beam energy, beam density, plasma density and modulation index and the desired radiation frequency can be tuned using plasma wiggler period and beam energy.

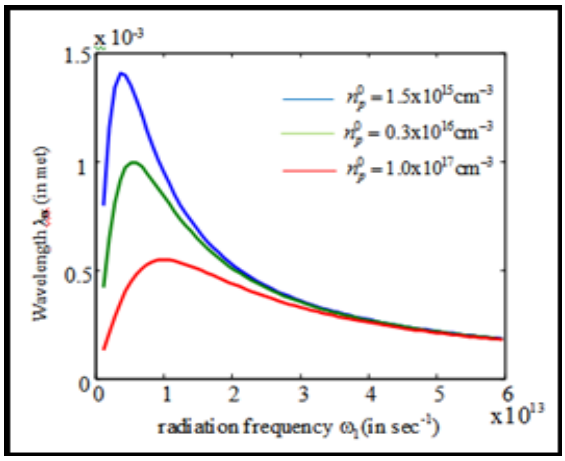


Figure 1. Figure showing that the periodic plasma wavelength decreases as plasma density increases and becomes sharper for THz radiation

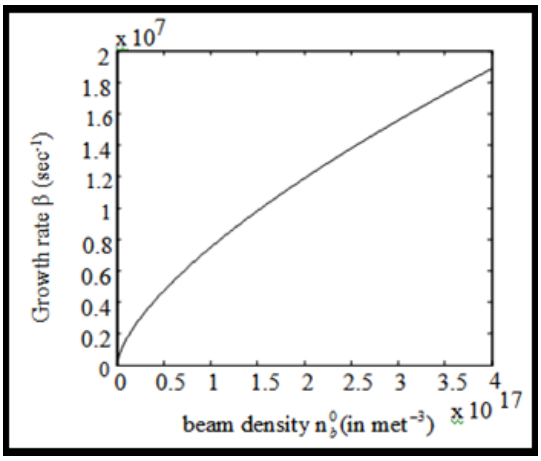


Figure 2. Variation of growth rate β of radiation wave with beam density n_b^o.

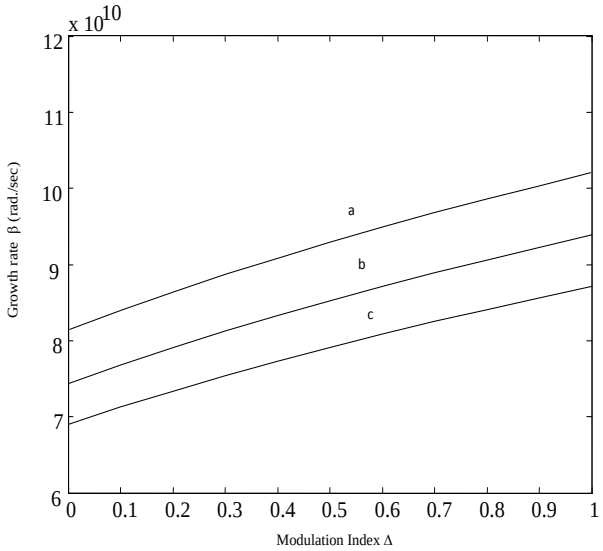


Figure 3. Variation of growth rate (β) with modulation index (Δ) at beam energy (a) E_b =1.0 MeV (b) E_b =2.0 MeV and (c) E_b =3.0 MeV

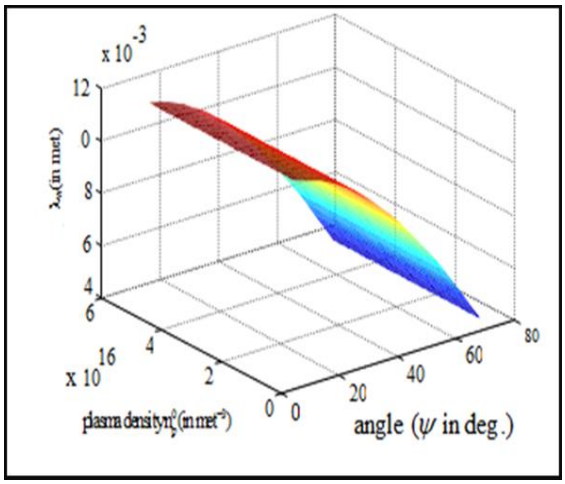


Figure 4. Variation of periodic plasma plasma wavelength λ_w with angle ψ and density n_p^o.

References

[1] Ferguson B and Zhang XC 2002 *Nat. Mater.* **1** 26
[2] Tonouchi M 2007 *Nat. Photon.* **1** 97
[3] Leemans W P, Tilborg J Van, Faure J, Geddes C G R, Toth C, Schroeder C B , Esarey E, Fubiani G and Dugan D 2004 *Phys. Plasmas* **11** 899
[4] HashimshonyD, Zigler A and Papadopoulos K 2001 *Phys. Rev. Lett.* **86** 2806

[5] Zhang X C, Bhu, Darrow J and Auston D 1990 *Appl. Phys. Lett.* **56** 1011
[6] Faist J, Capasso F, Sivco D L, Sirtori C, Hutchinson AL and Cho AY 1994 *Science* **264** 553
[7] Joshi C, Katsouleas T, Dawson J M, Yan YT and Slater J M1987 *IEEE Trans. J. Quant. Elect*. **23** 1571
[8] Miano R G and Vaccaro V G 1990 *Phys. Scr*. **30** 192
[9] Namiot V A and Shchurova LY 2012 *Phys. Lett.* A **375** 2759
[10] Zhang S C 2009 *Phys. Plasmas* **16** 093107

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Terahertz radiation generation by using modulated electron beam with plasma wave wiggler

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 J. Phys.: Conf. Ser. 836 012005

(http://iopscience.iop.org/1742-6596/836/1/012005)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 14.139.251.107
This content was downloaded on 06/06/2017 at 11:19

Please note that terms and conditions apply.

You may also be interested in:

Terahertz pulses in a periodic plasma channel via modulated electron beam
Pratibha Malik, Suresh C Sharma and Rinku Sharma

Terahertz radiation generation by the beating of two cross focused Gaussian laser beams in axially magnetized plasma
Saba Hussain, Ram Kishor Singh and R. P. Sharma

Influence of the Beam Self-Fields on the Dispersion Characteristics of EM Waves in a Dielectric Waveguide Filled with Plasma
Xiao Jixiong, Chen Shixiu, Tian Wei et al.

Low-frequency wiggler modes in the free-electron laser with a dusty magnetoplasma medium
S Jafari

Free-electron laser with a plasma wave wiggler propagating through a magnetized plasma channel
S Jafari, F Jafarinia and H Mehdian

Radiation from a quasi-periodically modulated electron beam

M. G. Kong and A. Vourdas
Electron trajectories and growth rates of the plasma wave pumped free-electron laser
S Jafari, F Jafarinia, M Nilkar et al.

Simultaneous Observation of RHEED Oscillation during GaAs MBE Growth with Modulated Electron Beam
Hideo Sugiura, Minoru Kawashima and Yoshiji Horikoshi

Intense relativistic electron beam ring (SPAC)
A. Mohri

IOP Conf. Series: Journal of Physics: Conf. Series **836** (2017) 012005 doi:10.1088/1742-6596/836/1/012005

Terahertz radiation generation by using modulated electron beam with plasma wave wiggler

J Panwar[a], S C Sharma and R Sharma

Department of Applied Physics, Delhi Technological University, Delhi-110 042, India

E-mail: [a)]panwar.jyots2010@gmail.com

Abstract. The generation of high power terahertz radiation has been studied by the interaction of density modulated relativistic electron beam (REB) with the plasma pump wave. The REB is pre bunched by the two laser beams which later translate to density modulation on passing through the drift space. The modulated beam enters the plasma medium and excites the electron plasma wave which leads to the oscillatory velocity of the electron beam. This transverse velocity couples with the modulated beam density and gives rise to nonlinear current density that leads to the emission of radiation of the order of terahertz. In results, the power of the radiation is evaluated and analysed. The power varies as square of the electron beam radius and field of the pump wave. The radiation thus obtained has numerous scopes in various fields.

1. Introduction

The generation of terahertz radiation using relativistic electron beams in free electron lasers has become an active area of research due to its wide applications in diverse fields including medical, security, spectroscopy, defence etc [1-2]. The emission of tunable terahertz radiation using pre bunched relativistic electron beam by two lasers has been studied by Kumar and Tripathi [3]. They noticed that the output power and amplitude of the radiation increases with the frequency. The frequency of the radiation was also found to be tuned by lasers wavelength. Plasma wave helps in obtaining higher output efficiency due to its short wavelength. Joshi et al.[4] have employed plasma density wave as a purely electric wiggler and studied the effect of plasma on the free electron radiation. They concluded that strong wiggler strength and short wiggler wavelength of plasma wave leads to emission of radiation at modest electron beam energy. Son et al.[5] have studied the emission of terahertz radiation using plasma wave as an undulator by forward Raman scattering and observed that the beam energy required reduces to a considerable amount and the energy spread reduces due to space charge effect. In this paper, we have studied the effect of relativistic electron beam and the pump wave on the emission of THz radiation, emitted by the interaction of density modulated relativistic electron beam with the space charge wave excited in the plasma medium.

2. Model

A relativistic density modulated electron beam is obtained following Kumar and Tripathi [3]. The two laser beams leads to the energy modulation of the REB by exerting a pondermotive force on it in the modulator.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

$v_m = -\lambda v_b \sin(\omega t_0 + \theta_b)\hat{z}$, where $\lambda = \frac{kP}{2\gamma^4}\frac{c^2}{v_b^2}\frac{e^2 A_a A_b^*}{m^2 c^2 \omega_a \omega_b}\frac{\sin(\theta_b)}{\theta_b}$, $\theta_b = \frac{(\omega - kv_b)P}{2v_b}$ is the phase angle of modulator, (ω, k) is the frequency and wave vector of modulated beam, P is the length of interaction region, v_b is the initial beam velocity, A_a, A_b and ω_a, ω_b are the electric field amplitudes and frequencies of the laser beams, respectively, γ is the relativistic gamma factor. The energy modulated beam on travelling through the drift space gets density modulated. $n_m = -2n_b J_1\left(\frac{\omega L\lambda}{v_b}\right)\cos(\omega t_3 - \psi)$, where n_b is the initial beam density.

The beam on interaction with the electrostatic pump wave acquires an oscillatory velocity $\overrightarrow{v_0} = \frac{e\overrightarrow{E_0}}{im\omega_0\gamma^3}$, where $\overrightarrow{E_0}$ is the pump electric field, ω_0 is the frequency of space charge pump wave. This velocity couples with the modualted beam density and leads to non-linear current density at phase (ω', k), where $\omega' = \omega + \omega_0$.

$$\vec{J}_{\omega',k} = \frac{2n_b e^2 A_0}{im\omega_0\gamma^3} J_1\left(\frac{\omega L\lambda}{v_b}\right) e^{-i(\omega' t - kz)} \tag{1}$$

This acts as an antenna and helps in evaluating the retarded vector potential. The power of the radiation is obtained using the poynting vector, calculated using the vector potential. The average power of the emitted THz radiation is evaluated as

$$4\pi r^2 \overrightarrow{S_{av}} = \frac{8\pi n_b^2 \omega'^2 e^4 A_0^2 r_e^4 l_e^2 J_1^2\left(\frac{\omega L\lambda}{v_b}\right)}{c^3 \varepsilon_0 m^2 \omega_0^2 \gamma^6}, \tag{2}$$

where ω' is the frequency of output radiation, l_e is the electron bunch length, r_e is the radius of beam. The normalized power is given as

$$\text{Normalized power} = \frac{2n_b^2 \omega'^2 e^4 r_e^4 l_e^2 J_1^2\left(\frac{\omega L\lambda}{v_b}\right)}{l_s^2 c^4 \varepsilon_0^2 m^2 \omega_0^2 \gamma^6}, \tag{3}$$

l_s is the length of space charge wave wiggler.

3. Results and discussions

The parameters used for calculating the power of the output radiation are: $v_b = 0.93c$, $n_b = 3\times10^{15} cm^{-3}$, $\gamma = 2.7$, $r_e = 50\mu m$, $l_e = 0.35cm$, $l_s = 0.4cm$. Using Eq. (3), we have plotted figure 1. The figure indicates that the normalized power of the output radiation first increases with the frequency, attains a maximum value and then decreases. This behaviour is observed due to the Bessel's function character of the density modulation. At a fixed value of frequency, the power is found to decrease with the wiggler length.

Figure 2 shows the same behaviour and for a particular value of frequency, the output power increases with the increase in beam density.

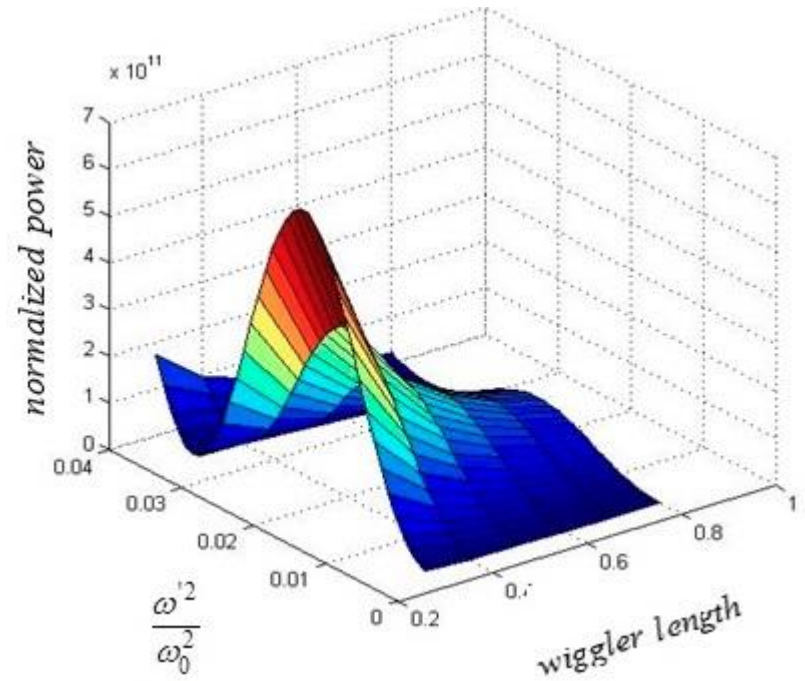


Figure 1. Figure showing the dependence of normalized power on the normalized frequency and wiggler length.

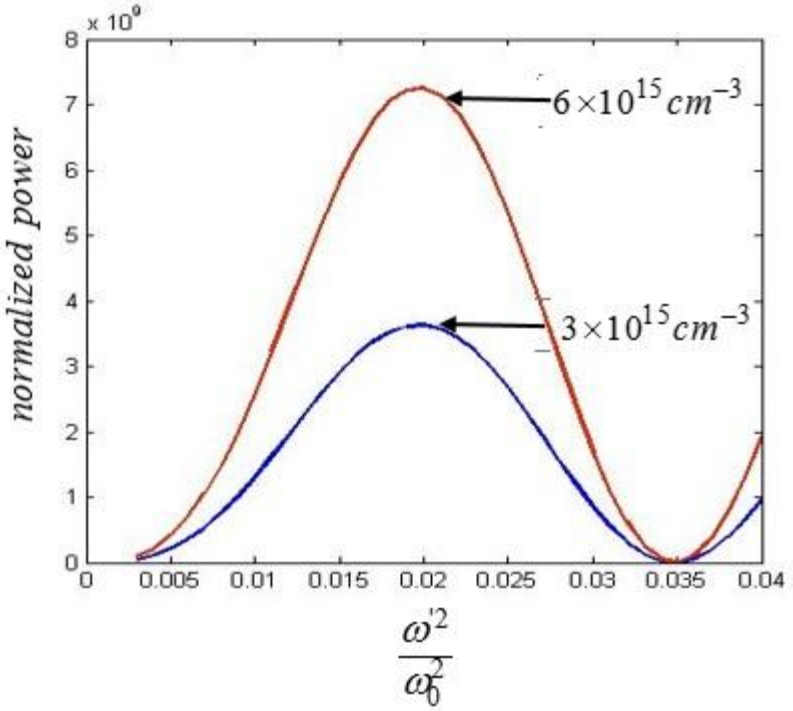


Figure 2. Figure showing the variation of normalized power with the normalized frequency for beam densities $n_b = 3\times10^{15}cm^{-3},\ 6\times10^{15}cm^{-3}$.

4. Conclusions

A pre bunched relativistic electron beam proves to be a great source for the generation of high power THz radiation. The power scales as square of the beam radius and beam density. It varies inversely as square of the wiggler length. The radiation can thus be tuned by varying the parameters of the REB and pump wave. The output radiation obtained can be used for practical applications where high power is required.

References

[1] Taylor Z D, Singh R S, Culjat M O, Suen J Y, Grundfest W S, Lee H and Brown E R 2008 *Opt. Lett.* **33** 1258
[2] Walther M, Fischer B M, Ortner A, Bitzer A, Thoman A and Helm H 2010 *Anal. Bioanal. Chem.* **397** 1009
[3] Kumar M and Tripathi V K 2012 *Phys. Plasmas* **19** 073109
[4] Joshi C, Katsouleas T, Dawson J M, Yan Y T and Slater J M 1987 *IEEE J Quant. Elec.* **23** 1571
[5] Son S, Moon S J and Park J 2012 *Opt. Lett.* **37** 5172

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 108C (2017) 887–896

Procedia Computer Science

International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

The versatility of an entropy inequality for the robust computation of convection dominated problems

Balaji Srinivasan[1] and Vivek Kumar[2]

[1] Indian Institute of Technology Madras, Chennai, India, sbalaji@iitm.ac.in
[2] Delhi Technological University, Delhi, India, vivekkumar.ag@gmail.com

Abstract

We present a discrete inequality that exhibits versatile uses in convection dominated problems. Much like the thermodynamic entropy inequality, the sign of this so-called discrete entropy production allows us to determine unphysical regions in the numerical solution without any *a-priori* knowledge of the solution. Further, the sign of the discrete production also functions as an excellent indicator for mesh adaptation in convection-diffusion and other singular perturbation problems. We also show preliminary results for how the operator can be used to derive robust schemes for convection dominated problems. All the above applications i.e. (a) Detecting unphysical numerical behavior, (b) mesh adaptation and (c) stabilization, are robust in that they are achieved without any *ad-hoc*, user introduced, parameters. We show a range of numerical results that exhibit the efficacy of the operator.

© 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

Keywords: Convection-Diffusion problems, Singular perturbation problems, Finite difference schemes, Burgers equation.

1 Introduction

It is well known that computations of flow problems are particularly difficult when they are dominated by convection. For non-turbulent, convection-dominated situations these problems are primarily:(1) Unphysical oscillations in *smooth* but high-gradient problems. These may be observed in steady or unsteady convection-diffusion problems near boundary or interior layers. (2) Unphysical oscillations in *non-smooth* "shock" problems. These occur around shocks in non-linear problems such as Burgers equation. (3) Completely unphysical *glitches* in non-linear problems. These are stable but unphysical structures which appear in the midst of seemingly stable solutions. Some examples are sonic glitches [1], the carbuncle phenomenon etc. [2]. The above are typically handled by either stabilization (upwinding) or choosing appropriate meshes for the problem [3]. While there are several stabilization schemes and meshing techniques that address these problems, these vary according to the problem and situation being chosen. There is also the invariable presence of some or the other ad-hoc parameter in the problem which

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science
10.1016/j.procs.2017.05.099

reduces robustness in the case of complex problems. So, there is need for coming up with a unified method for addressing stabilization and meshing.

Another important unaddressed question for numerical solutions of convection dominated problems is if one can actually determine that a particular solution is unphysical without *a-priori* knowledge of the solution. Addressing this question becomes particularly important in the case of novel problems or situations where one has no clear idea of the actual solution – a situation that occurs fairly commonly in practical computations. In what follows, we show how employing a simple entropy principle adapted from the study of hyperbolic conservation laws can address the above issues in a simple and unified way. The central idea behind the principle is as follows – *Just as the second law of thermodynamics separates unphysical solutions from physical ones by finding the sign of an entropy variable, it is possible to identify unphysical portions of a numerical solution by finding the sign of an appropriately defined discrete entropy production.*

Pioneering work on the employment of an entropy principle for numerical discretizations was done by Tadmor [4,5]. Tadmor characterized the minimum amount of numerical viscosity required for obtaining discretizations of hyperbolic conservation laws that would converge to the correct weak solution. Merriam [6] further adapted this discrete entropy principle to derive numerical schemes that satisfied the entropy principle on a cell by cell rather than global basis. Recently, Guermond et. al. [7] utilized an entropy inequality to derive stabilized, high-order finite element schemes for hyperbolic conservation laws.

All of the above apply an entropy principle to the computation of hyperbolic conservation laws. Further, the primary purpose in these has been to compute the correct weak solution and to calculate flow in non-smooth regions with minimal oscillations. To the best of our knowledge, there has been no work on deriving an equivalent principle for problems other than hyperbolic conservation laws, such as the convection-diffusion equation. Furthermore, it has not been investigated if the same idea could be applied to capturing boundary layers accurately in smooth flows.

In what follows, we demonstrate different uses for a unified entropy production principle for convection dominated problems. We show that the same principle may be applied for (a) Detecting unphysical numerical behavior without *a priori* knowledge of the solution, (b) To derive stabilization methods and (c) For deriving mesh adaptation schemes for singularly perturbed problems.

2 Entropy production, detection and stabilization of unphysical behavior

2.1 Unphysical behavior in the steady state convection-diffusion problem

Numerical solutions of convection dominated problems show various numerical artefacts that are particularly visible in high gradient regions of the solution. The simplest example of this is observed in the solution of the steady state convection-diffusion problem.

$$\frac{du}{dx} = \varepsilon \frac{d^2u}{dx^2} \tag{1}$$

with specified boundary conditions. This problem, though simple, has attracted much analysis in the literature as it represents a simple but powerful prototype of the type of behavior seen

in boundary layer problems with the full Navier-Stokes equations. A host of explanations and solutions have been offered in the literature for this problem (see [8] for a comprehensive review). However, all of them invoke some *a priori* knowledge of the monotonic behavior of the solution. We provide a new explanation for this numerical behavior which does not depend on any *a priori* knowledge of solution behavior and also extends easily to other problems. For this, we recall the entropy function as used in the study of hyperbolic conservation laws [9]. Given a scalar hyperbolic conservation law of the form $u_t + f(u)_x = 0$, one can define a corresponding entropy variable $S(u)$, entropy flux $F(u)$ pair provided they satisfy $S_{uu} > 0$, and $F_u = S_u f_u$. It follows immediately that for a smooth solution $S_t + F(u)_x = 0$. However, for solutions with discontinuities, it can be shown that only the correct weak solution satisfies the *entropy inequality* $S_t + F(u)_x \leq 0$. Merriam [6] first observed (without proof) that the violation of a discretized version of the above entropy inequality, correlates well with the presence of shock oscillations for scalar hyperbolic conservation laws. We now derive a similar equation for the (non-hyperbolic) convection-diffusion equation and prove that the satisfying or violation of an entropy principle correlates exactly with the local presence or absence of oscillations even for smooth solutions.

Multiplying equation (1) with $2u$, one obtains $2u\frac{du}{dx} = \varepsilon\left(2u\frac{d^2u}{dx^2}\right)$. On expanding the RHS (the diffusive term), this can be simplified and rearranged to obtain

$$\frac{d(u^2)}{dx} - \varepsilon\frac{d^2(u^2)}{dx^2} = -2\varepsilon\left(\frac{du}{dx}\right)^2 \tag{2}$$

The negative term on the right hand side of this auxiliary equation is the scalar equivalent of the dissipation term in the kinetic energy equation for the incompressible Navier-Stokes equations, which is known to contribute to the entropy production in the internal energy equation. Therefore, by analogy to the Navier-Stokes equations and the similar equation in the case of hyperbolic conservation laws, we term this auxiliary equation the "entropy production" equation. We, therefore, obtain the following entropy production inequality that all (smooth) solutions of equation (1) must necessarily satisfy,

$$\frac{dS}{dx} - \varepsilon\frac{d^2S}{dx^2} \leq 0 \tag{3}$$

where $S(u) = u^2$ is the entropy variable. The key observation here is that not all discretizations of (1) need satisfy a discrete version of (3). *We term a discrete solution $u(x_i)$ of (1) as a "discrete entropy satisfying" solution if it obeys an appropriately chosen, natural discretization of (3).* By a natural discretization we mean that each of the derivative operators in the entropy equation is discretized identically to the corresponding operator in the discretization in (1). For instance, for the 2^{nd} order central difference scheme on a uniform grid, (1) is discretized as

$$\frac{u_{i+1} - u_{i-1}}{2h} = \varepsilon\left(\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}\right) \tag{4}$$

Consequently, (3) is naturally discretized as

$$P_i = \frac{S_{i+1} - S_{i-1}}{2h} - \varepsilon\left(\frac{S_{i+1} - 2S_i + S_{i-1}}{h^2}\right) \leq 0 \tag{5}$$

The above equation is the discrete entropy inequality corresponding to the central discretization of the convection diffusion equation. Similarly, the upwind discretization of convective terms

in (1)

$$\frac{u_i - u_{i-1}}{h} = \varepsilon \left(\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \right) \tag{6}$$

would have the corresponding discrete entropy inequality

$$P_i = \frac{S_i - S_{i-1}}{h} - \varepsilon \left(\frac{S_{i+1} - 2S_i + S_{i-1}}{h^2} \right) \leq 0 \tag{7}$$

The utility of the above discrete inequalities is due to the following properties which they satisfy:(1) Solutions of (4) exhibit unphysical oscillations if and only if the discrete entropy inequality (5) is violated. (2) The correct threshold value of $\epsilon (= h/2)$, above which the central difference solution (4) is monotonic is automatically obtained by enforcing (5). (3) The upwind discretization of convective terms in (1) results in the discrete entropy inequality to be always satisfied.

Proofs We offer concise proofs of the above properties now. Property 1 can be proved by multiplying (4) by $u_{i+1} + u_{i-1}$ and subtracting from (5). Rearrrangement gives

$$P_i^{central} = -2\varepsilon \left(\frac{u_{i+1} - u_i}{h} \right) \left(\frac{u_i - u_{i-1}}{h} \right) \tag{8}$$

Now, the discrete entropy inequality (5) is satisfied at any grid location x_i if $u_{i+1} - u_i$ and $u_i - u_{i-1}$ have the same sign i.e. the solution is monotonic in $[x_{i-1}, x_{i+1}]$. Similarly, the presence of oscillations around a grid location automatically indicates that $u_{i+1} - u_i$ and $u_i - u_{i-1}$ have different signs, immediately violating the inequality (5), completing the proof.

Property 3 can be similarly proved by multiplying the upwind discretization (6) by $u_i + u_{i-1}$ and subtracting from (7). Rearrangement gives $P_i^{upwind} = -2\varepsilon \left(\frac{u_i - u_{i-1}}{h} \right)^2$. Hence the discrete entropy inequality is satisfied for all positive values of ϵ for an upwind scheme, completing the proof.

Finally, Property 2 can be proved by letting $F = \frac{u_{i+1} - u_i}{h}$ and $B = \frac{u_i - u_{i-1}}{h}$. Then, (4) satisfies
$\frac{F+B}{2} = \varepsilon \frac{F-B}{h}$. Writing B in terms of F and substituting in (8) gives,$P_i = 2\epsilon \left(\frac{1/2 - \varepsilon/h}{1/2 + \varepsilon/h} \right) F^2$
We immediately conclude that $|\varepsilon| \geq h/2$ for satisfying the entropy inequality, completing the proof. We note the following:

- The discrete entropy production operator P_i is simply calculated *a posteriori*. Wherever the inequality is violated, one immediately knows that the corresponding region of the solution is untrustworthy. We require no *a priori* knowledge of what the solution looks like. This allows to extend this application to other equations below.

- The amount of violation of the inequality (how positive P_i is) is directly related to how oscillatory the solution is. Figure 1 confirms this through solutions of the convection-diffusion equation (1) for two different ϵ along with the corresponding values of the entropy production. One can observe that a local violation of the entropy inequality matches perfectly with unphysical behavior.

- This derivation of the threshold value seems to be the first one in the literature which is done completely independent of the requirement of non-oscillatory behavior of the solution. We re-emphasize that all we have required of the solution is that it obey the discrete entropy inequality and the correct threshold value immediately follows.

- We emphasize the remarkable fact that an operator originally derived for identifying the physically correct *discontinuous* solution in hyperbolic conservation laws is able to automatically separate out the incorrect numerical behavior within a completely *smooth* solution.

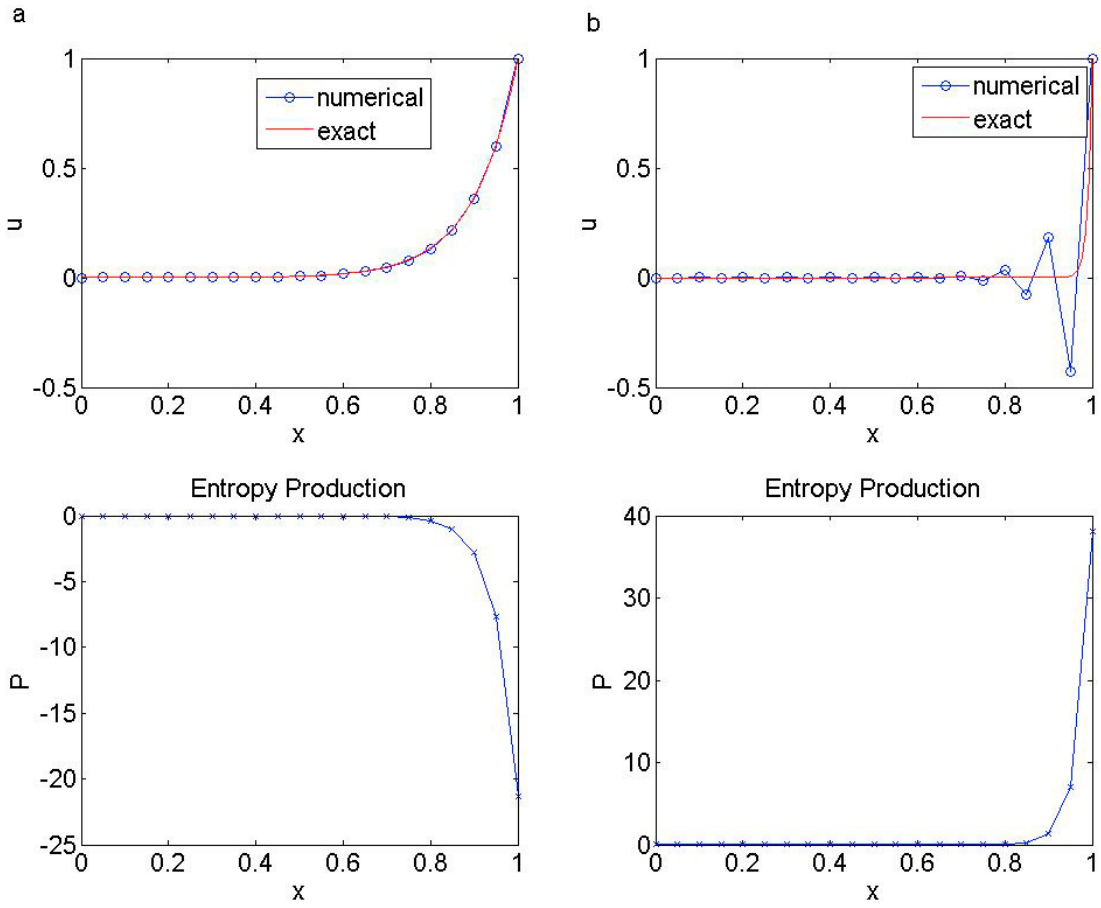


Figure 1: Finite-difference solutions for convection diffusion equation for N = 20 points in the domain. Note the correspondence between positive entropy production and unphysical oscillations (a) Well resolved case with $\epsilon = 0.1$ (b) Under-resolved case with $\epsilon = 0.01$

2.2 Extensions to other convection-dominated equations

2.2.1 Unsteady advection-diffusion

The above principle may be applied to the unsteady advection-diffusion equation $u_t + au_x = \varepsilon u_{xx}$ with the corresponding entropy inequality $P = S_t + aS_x - \varepsilon S_{xx} = -\varepsilon(u_x)^2 \leq 0$. As before, discretization of this inequality is done by exactly repeating the discretization of the original equation operator by operator. For instance, for Euler time-stepping and central derivatives, Figure 2 shows a temporally evolving "square wave" for two different values of the coefficients. As can be seen, for the convection-dominant case the advent of oscillations is correlated with

the violation of the entropy inequality. Further, even within the oscillatory solution, the smooth sub-regions are correctly identified as entropy satisfying regions.

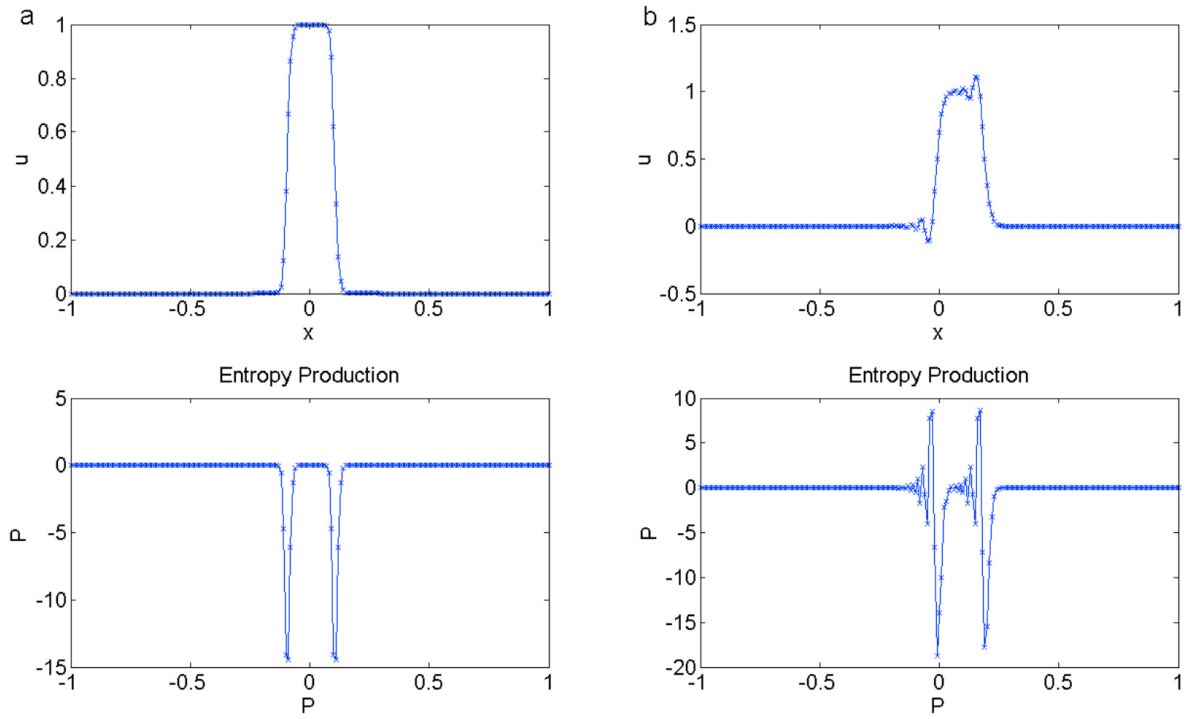


Figure 2: Advection-Diffusion equation 2^{nd} order central difference simulation with 200 points. (a) Non-oscillatory solution for $a = 1$, $\epsilon = 0.01$ (b) Stable but oscillatory solution for $a = 10, \epsilon = 0.01$. Notice that even within the front and back of the wave the smooth regions are correctly identified as being entropy-satisfying.

2.2.2 Scalar singular perturbation problems for convection-dominated problems

The same principle can be extended to any general scalar singular perturbation problem of the form $-\varepsilon u"(x) - a(x)u'(x) + b(x)u(x) = f$ with the corresponding entropy inequality $P = -\varepsilon S''(x) - a(x)S'(x) - 2uf + 2b(x)u(x) - 2\varepsilon(u')^2 \leq 0$. The numerical results for this are included in Section 3 for general, non-uniform grids

2.2.3 Nonlinear case – Viscous Burgers equation

The viscous Burgers equation is given by $u_t + \left(\frac{u^2}{2}\right)_x = \varepsilon u_{xx}$ with the corresponding entropy inequality is derived to be $P = S_t + \left(\frac{2u^3}{3}\right)_x - \varepsilon S_{xx} = -2\varepsilon(u_x)^2 \leq 0$. To test the efficacy of our entropy inequality, we consider the following two classic Riemann problems

Case 1: *Under-resolved shock.* $(u_L > u_R)$. We present a steady state Burgers shock for which the exact solution is also known. As seen in Figure 3a, when there are no oscillations in the well-resolved case (i.e. ϵ is above a certain threshold value), the entropy inequality is satisfied uniformly. Further, as seen in Figure 3b, one obtains oscillations near the shock when

the shock is under-resolved and no oscillation in case it is well resolved. One can see that the entropy inequality is violated precisely wherever the unphysical oscillation occurs.

Case 2: *Under-resolved expansion fan* ($u_L < u_R$). Improperly resolved expansion fans lead to a phenomenon known as the sonic glitch [1]. This is an $O(h)$ "expansion shock" at the sonic point right in the middle of a smooth region and is famously exhibited by even entropy-satisfying methods such the Godunov method [1]. The problem, of course, is that a method may satisfy the entropy condition in the limit of infinitesimal grid sizes but violate it in the discrete case.

Figure 4 shows an under-resolved expansion with initial conditions being $u_L = -1$ and $u_R = 1$. We have chosen $\epsilon = 0.5$ for this problem as it clearly exhibits the sonic glitch. It can be seen that has the entropy inequality is violated right around the sonic glitch while the inequality is satisfied at other smooth regions. To our knowledge, this is the first time in the literature that this violation of discrete entropy around the sonic point has been shown.

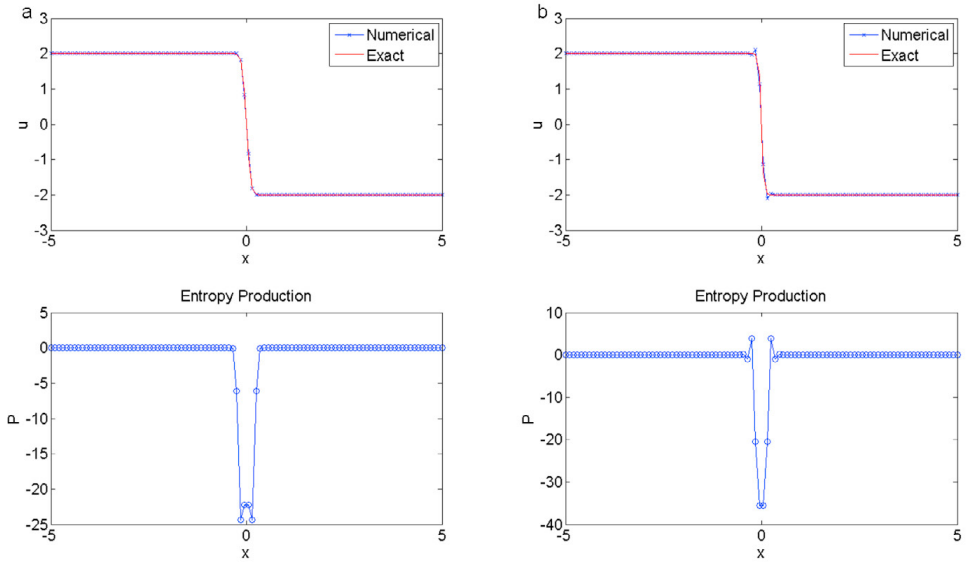


Figure 3: Entropy production for a steady state Burgers shock with $u_L = 2$ and $u_R = -2$, with 100 equally spaced mesh points (a) Well resolved solution with $\epsilon = 0.1$ (b) Oscillatory solution for $\epsilon = 0.05$.

2.2.4 An attempt to stabilize the sonic glitch

We now show that an adapted entropy viscosity method can be used to fix the sonic glitch. Based on our conjecture that the sonic-glitch is essentially due to local entropy violation, we adapt the method of Guermond et. al. [7] which works well for shocks, for this expansion fan case. We have one important difference from the method in [7]. In order to enforce Galilean invariance, we calculate the entropy based on a local shift of velocity i.e. for every grid point x_i , the entropy production is calculated using the triad of velocities $[u_{i-1} - u_i, 0, u_{i+1} - u_i]$. Correspondingly, we use the entropy function $S = (u - u_i)^2$ locally. It can be easily seen that the corresponding entropy flux is therefore modified to $F = \frac{2}{3}\left(u - \frac{3u_i}{2}\right)u^2$. The corresponding entropy equation is $P = S_t + F_x - \varepsilon S_{xx} = -2\varepsilon(u_x)^2 \leq 0$. As can be seen in Figure 5, the results we obtain are exceptionally good with our choice. A comparison with the solution obtained

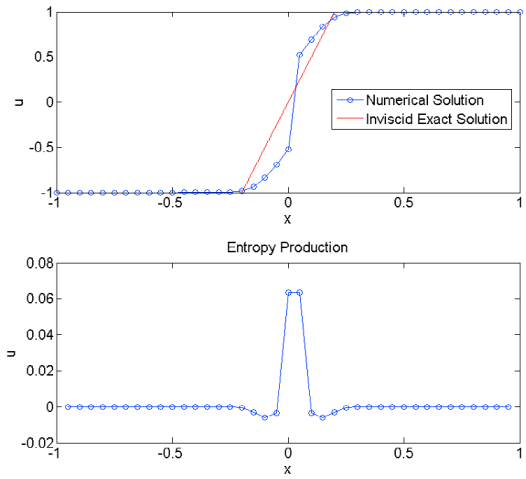


Figure 4: Entropy production for a expansion fan with $u_L = -1$ and $u_R = 1$, and 40 equally spaced mesh points. An under-resolved solution with $\epsilon = 0.05$ at a time of 0.2 is shown. Entropy violations are clearly related to the sonic glitch.

with the standard Godunov method [1] for the convective terms shows how much more effective the current method is.

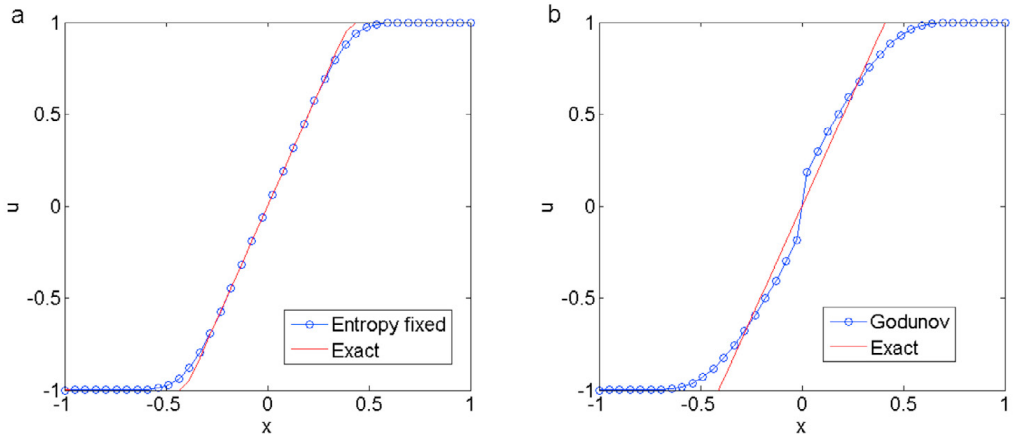


Figure 5: Solution at t = 0.2 for an expansion fan. (a) The entropy fixed solution which does not show the sonic glitch (b) The Godunov solution which shows a clear sonic glitch around the sonic point.

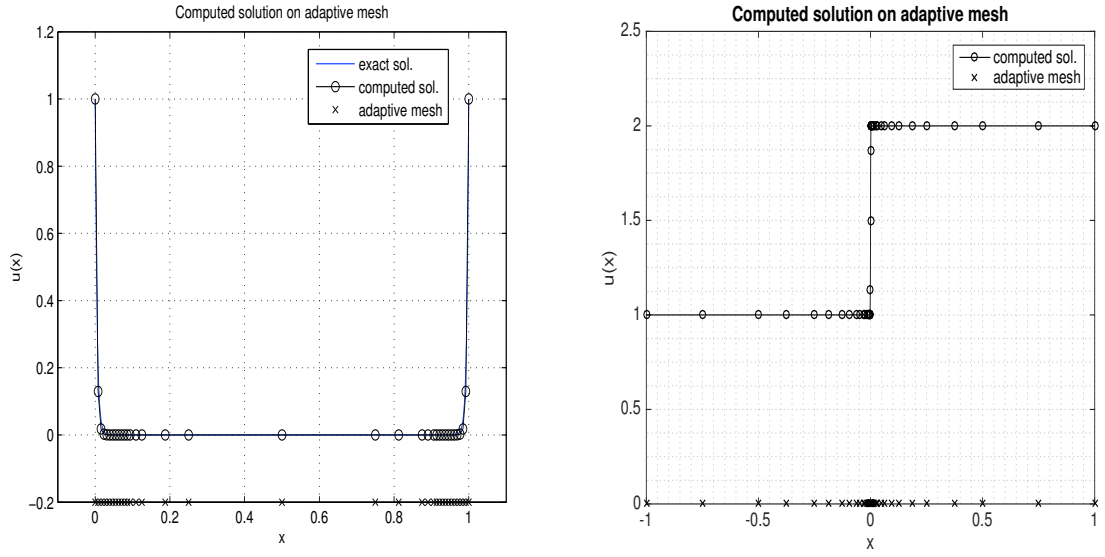
3 Adaptive meshes for singular perturbation problems

Since the entropy inequality detects regions of unphysical behavior, it can be adapted, almost unmodified, for adaptive mesh generation for singular perturbation problems. The general procedure is to compute the *a posteriori* entropy residuals for a given mesh and adapt around the region of maxium entropy production. Unlike other standard meshes for singularly perturbed

problems, such as Shishkin and Bakhvalov meshes [3] our method requires no prior knowledge of the where the boundary layer region lies and requires no arbitrary parameters. We refer the reader to [11] for the details of the implementation of the above strategy as well as detailed numerical results for a variety of singular perturbation problems. Here, we just present a couple of results that exhibit the versatility of the method.
Turning Point Problem – We consider a SPP with a turning point [11] $-\varepsilon u'' + 2(2x-1) + 4u = 0;\ x \in (0,1)$, with the BCs $u(1) = 1$; $u(0) = 1$. This problem has a turning point, i.e. a sign change of the convective term at $x = 1/2$. Consequently, there are two boundary layers in this problem. Figure 6a shows the final solutions for this problem with $\epsilon = 10^{-5}$ and $N = 55$.

Interior Layer Problem – As another example, we consider a SPP with an interior layer point $-\varepsilon u'' - xu' = 0;\ x \in (-1,1)$ with the BCs $u(-1) = 1$; $u(1) = 2$. This problem has an interior layer, i.e. a "layer" in the inside of the domain instead of the boundaries at $x = 0$.Figure 6b shows the final solutions for this problem for $\epsilon = 10^{-6}$ and $N = 45$.

In both cases, we started with *just one point within the domain* excluding the boundaries and provided no input on where the sharp gradients were. We notice once again the excellent results that are obtained without any input on the nature of the problem. Apart from these cases, we have performed several others with equally excellent results. We know of no other conventional method that works with such minimal inputs on both the number of points and nature of the problem. This method for mesh adaptation therefore seems to us to be quite promising.



(a) Turning point, $\epsilon = 10^{-2}$ with $N = 35$ adaptive mesh points

(b) Interior layer,$\epsilon = 10^{-6}$ with $N = 45$ adaptive mesh points

Figure 6: Adaptive Meshes for singular perturbation problems. The mesh automatically adapts to both cases with just N = 3 points to start with.

4 Conclusions and future directions

We have presented in this paper a method to utilize a suitably defined discrete entropy production operator and a corresponding inequality for convection dominated problems. For the case of the convection-diffusion equation we were able to prove that the violation of the inequality is directly related to the advent of unphysical behavior even in smooth regions. In other convection-dominated cases we presented numerical results that also indicated the ability of this inequality to identify regions with unphysical behavior such as oscillations and glitches. Preliminary results for how to utilize this method to fix the identified problems in case of the sonic glitch were also presented. Finally, it was also shown that the same operator could be used in SPPs to generate excellent adapted meshes for a range of problems. For utilization in practical problems, the method needs several extensions. Firstly, the method needs to be tested on high order finite difference schemes. Also, extensions for finite volume and finite element methods would allow the method to be used for problems with complex geometries. Our preliminary results for these appear promising [11]. Finally, we wish to extend the method to become capable of handling multiple variables (instead of the scalar cases we have presently done). We shall report on our progress on these questions in future publications.

Acknowledgements

The second author acknowledges the research project No. $2/48(6)/2016/NBHM(R.P.)/R\&D\ II/15455$.

References

[1] Tang, H. (2005). On the sonic point glitch. Journal of Computational Physics, 202(2), 507-532.

[2] Dumbser, M., Moschetta, J. M., & Gressier, J. (2004). A matrix stability analysis of the carbuncle phenomenon. Journal of Computational Physics, 197(2), 647-670.

[3] Lin, T., & Stynes, M. (2001). Numerical methods on Shishkin meshes for linear convection–diffusion problems. Computer methods in applied mechanics and engineering, 190(28), 3527-3542.

[4] Tadmor, E. (1984). Numerical viscosity and the entropy condition for conservative difference schemes. Mathematics of Computation, 43(168), 369-381.

[5] Tadmor, E. (1987). The numerical viscosity of entropy stable schemes for systems of conservation laws. I. Mathematics of Computation, 49(179), 91-103.

[6] Merriam, M. L. (1989). An entropy-based approach to nonlinear stability. NASA Technical Memorandum, 101086(64), 1.

[7] Guermond, J. L., Pasquetti, R., & Popov, B. (2011). Entropy viscosity method for nonlinear conservation laws. Journal of Computational Physics, 230(11), 4248-4267.

[8] Oate, E., & Manzan, M. (2000). Stabilization techniques for finite element analysis of convection-diffusion problems. Developments in Heat Transfer, 7, 71-118.

[9] Lax, P. D. (1973). Hyperbolic systems of conservation laws and the mathematical theory of shock waves (Vol. 11). SIAM.

[10] Puppo, G. (2004). Numerical entropy production for central schemes. SIAM Journal on Scientific Computing, 25(4), 1382-1415.

[11] Kumar, V., & Srinivasan B. (2015). An adaptive mesh strategy for singularly perturbed convection diffusion problems. Applied Mathematical Modelling, 39(7), 2081-2091

Nuclear Instruments and Methods in Physics Research B xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research B

journal homepage: www.elsevier.com/locate/nimb

Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique

Gaurav Sharma [a,b], Nitin K. Puri [a,*], T. Nandi [b,*]

[a] Dept. of Applied Physics, Delhi Technological University, New Delhi, Delhi 110042, India
[b] Inter University Accelerator Centre, New Delhi 110067, India

ARTICLE INFO

Article history:
Received 30 November 2016
Received in revised form 26 May 2017
Accepted 29 May 2017
Available online xxxx

Keywords:
X-ray spectroscopy
X-ray filtration
Pile-up
Resoluion

ABSTRACT

The X-ray absorption spectroscopy technique has been discussed for resolving the low-intensity peaks obscured in the neighborhood of the closely spaced peaks and eliminating the pile-up effect, simultaneously. A theoretical comparison of the absorption technique with the available pile-up rejecters reveals better efficiency for the detection of the higher energy X-ray lines. The technique does not change the fundamental characteristics of the peaks and therefore can be employed to measure the lifetimes of the corresponding states and absolute cross sections. The lifetime of a state calculated with this technique is found to be the same as that without attenuating the intensity using any absorber.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Spectral resolution of detectors has been a crucial issue for the detection of closely lying peaks. With the development of high resolution detectors, the efforts have been made further to improve the resolution. The maximum resolution of a detector can be achieved by minimizing the overlap of the two peaks. This overlap can be reduced either by increasing the separation between their centroids or by reducing the full width half maxima (FWHM) of the peaks; otherwise another detector with higher resolving power will be required. Nevertheless, for any detector the experimental limitation may restrict the detector from resolving the two closely spaced peaks even though they lie within the resolving power of the detector. Such problem may appear in the experiments if the intensity of the two closely spaced peaks differs a lot from each other and leads to the blending, and sometimes the small peak disappears depending on the intensity ratio of the peaks. One possible method of resolving these two closely lying lines is by reducing the intensity of the intense peak which can be done either by X-ray absorption near-edge spectroscopy (XANES) [1,2] or decaying its intensity with time by shifting the detector away from the source [3]. In the former method, a specific absorber of a particular K edge energy will be required while the later can only be applied if the lifetime of the intense peak is much shorter than the weak transition.

Further, the pile-up [4] phenomenon is initiated due to a high count photon rate in the measurement, which leads to the generation of an artificial peak at twice the photon energy of the intense peak and a background is introduced between the region of the original peaks and artificial peak. If the original peaks present in the pile-up affected region, has low intensity; they will not be observed in the spectrum. The processes such as two-electron-one-photon [5–8] and multiply excited states having a low transition probability often lie near or in the pile-up effected region. One important point is to note that the relative yield for these processes to the singly excited processes is of the order of 10^{-4} for Ar atom [9]. This ratio improves for lighter atoms and becomes worse on the heavier atoms [5]. Presently, several pile-up rejection techniques including pulse shape discrimination[10], pile-up rejection circuits, reduction of the counting rates, etc [4,11,12] are known; but due to their dependence on dead time of the detector, they cannot eliminate the pile-up completely [4]. Apart from this, all these methods reduce the counting rate in the same ratio for all the peaks and so with the low intense peaks, which leads to further decrease in their count rate making their detection more difficult.

To resolve the above mentioned issues simultaneously, the X-ray absorption technique can be used that not only resolves the low intense peaks from their intense neighborhood but also

* Corresponding authors.
E-mail addresses: gaurav.kmc891101@gmail.com (G. Sharma), nitin.phy@dce.edu (N.K. Puri), nanditapan@gmail.com (T. Nandi).

http://dx.doi.org/10.1016/j.nimb.2017.05.062
0168-583X/© 2017 Elsevier B.V. All rights reserved.

Please cite this article in press as: G. Sharma et al., Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.062

minimizes the pile-up problems. The efficiency of this technique for detection of high energy photons existing near the pile-up region is found to be better than the electronic methods mentioned above [4,11,12]. This technique uses the concept of the strong variation of attenuation coefficient of absorber with the incident photon energy [1]. This technique is nondestructive and does not change the fundamental characteristics of the peaks. Therefore, the lifetimes of the states and absolute cross sections can be measured using this technique.

2. Disappearance of the peaks

Theoretically, a detector can efficiently resolve two peaks if they are separated by the sum of half width at half maxima (HWHM) of the two peaks. In spite of this condition being met by a detector, one experimental limitation can often be encountered if one of the peaks is much more intense than the others. Then the low-intensity peaks tend to be veiled in the broad base of the intense peak and consequently they cannot be resolved (Fig. 1). This practical problem cannot be solved until we modify the intensities of the peaks. Unfortunately, during the experiments we do not have any control over the intensities of the peaks. Therefore, many low intense peaks remain unidentified even though they are within the resolving power of the detector.

The disappearance of individual identity depends on the characteristics of both the peaks and the separation between them. For showing such phenomena we have simulated two closely separated Gaussian curves and varied their intensities (or amplitude), as shown in Fig. 1. For convenience, the peaks of same FWHM but with different separations are used, which are shown in Fig. 1. Fig. 1a(i) shows a well-resolved nature of the two peaks which keep on disappearing as the intensity of one of these peaks is enhanced, as shown in Fig. 1 a(ii) and (iii). Fig. 1a(iii) shows that as the amplitude ratio >10:1 the low-intensity peak merges to the intense peak. This ratio R, required for disappearance, varies with the characteristics and separation of the peaks. If the separation between the peaks is increased, the value of R required for disappearing the small peak will also increase and thus so with the resolution of the peaks. The value of R may vary from a few tens to a few orders of ten depending upon the nature of the two nearby peaks; a few examples in Fig. 1 showing the variations in R. In this section we have given a formulation for the calculation of the intensity required for disappearance of the low-intensity peaks.

As the transition probability is directly proportional to the intensity, the formulation has been derived in terms of the intensity as the characteristics of the peaks. The ratio, R, required for the disappearance of the low intense peak can be given by:

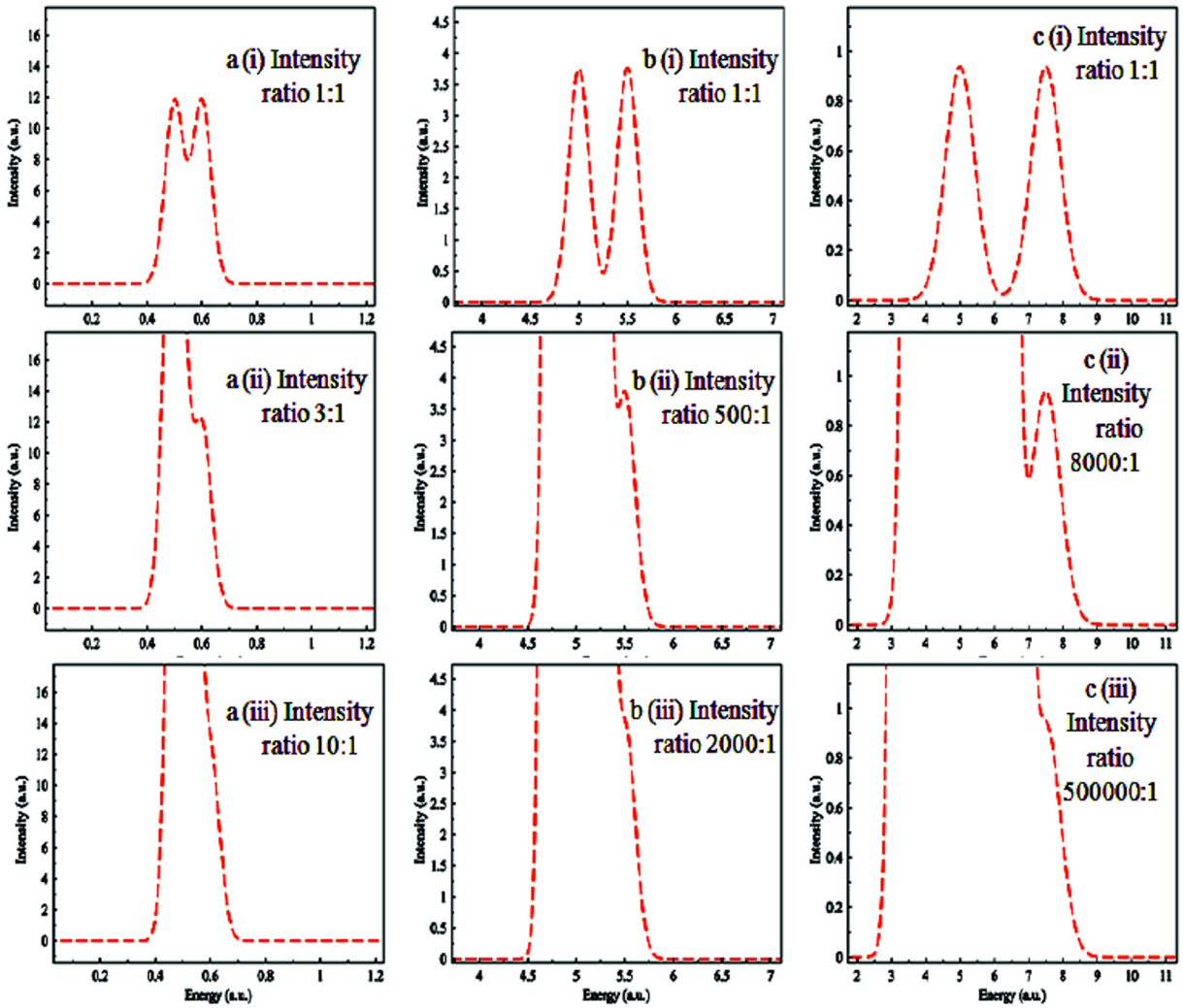


Fig. 1. A comparison of resolution on the width of the peaks and closeness to each other is shown. Fig. a(i)–c(i) show two completely resolved peaks and their dependence on the separation between their centroid. Fig. a(ii)–c(ii) and a(iii)–c(iii) shows the partially and completely disappeared peaks respectively. The peak disappearance for different value of ratio (R) shows the role of peaks characteristics in the resolution.

Please cite this article in press as: G. Sharma et al., Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.062

$$R = \left\{ \left(\frac{100}{P}\right) * \exp\left[\frac{(-ln2) * \left[(x_{c1} - x_{c2}) + \sqrt{\ln\left(\frac{P}{100}\right) * \frac{w_1^2}{-\ln(2)}} \right]^2}{w_2^2} \right] \right\}^{-1} \quad (1)$$

where, x_{c1} and x_{c2} are the centroids of the peaks and w_1 and w_2 are the HWHM of the peaks. P is the percentage value of intensity at the intersection point of two Gaussian curves to the maxima of the low intense peak. The value P decides the level of disappearance of the low-intensity peak in the shades of the intense peak. A comparison of the disappearance of low-intensity peak with P values is shown in Fig. 2. Simulation of the peaks in Fig. 2 suggests the effective disappearance for P value from 75 onwards and can have a maximum value for 100. Choosing a maximum P value, Fig. 2(f), we can obtain the intensity or the amplitude required for the complete disappearance of the low-intensity peak. This disappearance with the P values is verified with many possible situations of the peaks.

The ratio R can also be rewritten in terms of area, given by:

$$R = \frac{A_2}{A_1} = \left(\frac{Fw_1}{Fw_2}\right) * \left\{ \left(\frac{100}{P}\right) * \exp\left[\frac{(-ln2) * \left[(x_{c1} - x_{c2}) + \sqrt{\ln\left(\frac{P}{100}\right) * \frac{w_1^2}{-\ln(2)}} \right]^2}{w_2^2} \right] \right\}^{-1} \quad (2)$$

Here (Fw_1, Fw_2) and (A_1, A_2) are the FWHMs and area of the two adjacent peaks, respectively. This Eq. (2) gives the required intensity or amplitude ratio for merging the peaks from FWHM and position of the two peaks. Therefore, the absorption required for resolving the two unresolved peaks can be calculated. Another aspect of this formulation is to decide whether a peak of certain FWHM may exist in the wings of the intense peak by applying the minimum disappearance condition ($P \sim 75$).

3. Technique

The x-rays are absorbed in the medium dependent upon the thickness (t), density (ρ), mass number (Z) and atomic mass (A) of the material according to the equation given below. Further, the absorption coefficient (μ) varies as E^{-3}, where E is the energy of incident photons.

$$I = I_0 \exp(-\mu t) \quad and \ \mu = \frac{\rho * Z^4}{A * E^3} \quad (3)$$

This strong and smooth dependence of absorption coefficient μ provides a valuable tool to alter the intensities disproportionately i.e. absorption of the x-rays will be different for the different energies. Eq. (3) suggests; for the same absorber, the lower energy peaks will be absorbed more strongly than the peaks with higher energies.

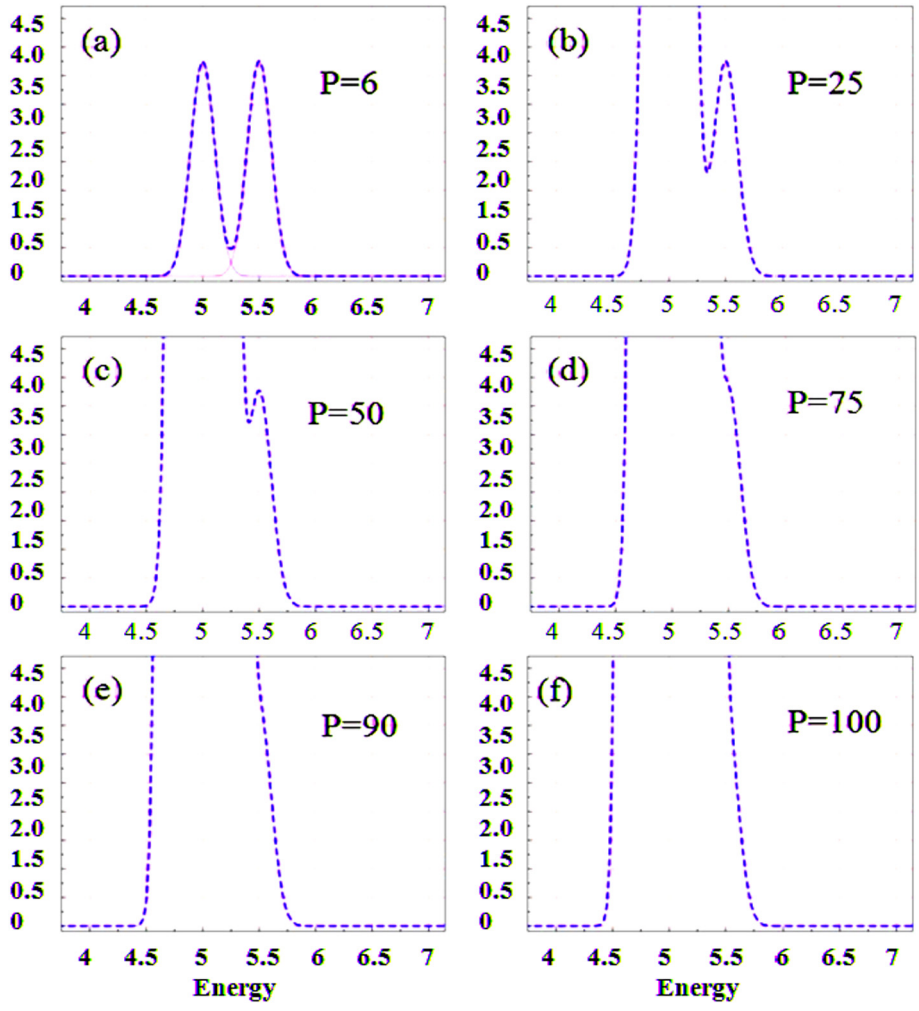


Fig. 2. Showing the effect of the P value on resolution of adjacent peaks. (a)–(c) showing the clear presence of second peak in the spectrum, from P value 75 onward the peak disappears (d)–(f).

Please cite this article in press as: G. Sharma et al., Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.062

4. Possible applications of absorber technique

4.1. Resolving peaks using absorber: technique to resolve

Now, as mentioned above that since the unification problem appears due to the much higher intensity of the nearby peaks, it can be avoided by decreasing the intensity of stronger peak. Using the presently available electronic techniques the intensities of the peaks can be altered in equal ratio by varying the detection time of the detector. Reducing the source strength also does not work for improving the resolution as all the peaks will be reduced proportionally. So, in order to achieve the spectral resolution, the intensities must be changed disproportionately to achieve nearly 1:1 ratio along with the resolution back to its optimum.

The proposed technique is discussed to resolve such problem by getting the desired intensity ratio using the benefit of energy dependence of the Lambert's law given in Eq. (3). So, from Fig. 1 we need to decrease the intensity I_1 to I_1' and I_2 to I_2' so for a particular thickness $I_1' \sim I_2'$. Calculations suggest that the absorber thickness required for the same amplitude in terms of energy of the corresponding photons is given by:

$$Thickness\ (t) = \left(\frac{\rho * Z^4}{A}\left(\frac{1}{E_1^3} - \frac{1}{E_2^3}\right)\right)^{-1} * \ln\left[\frac{I_{01} * w_1}{I_{02} * w_2}\right] \quad (4)$$

E_i, I_{0i} and w_i are the energies, intensities and HWHM of the two peaks to be resolved, respectively. Thus the unresolved peaks due to unequal amplitudes of the peaks will be resolved after having attenuation in the intensity by an absorber of thickness t, given by Eq. (4).

4.2. Pile-up rejection

Pile-up [4] is an affect on the detector and related electronics occurring in case of the high count rate processes. If the incident photon rate is more than the detection system can handle, then it creates a false peak at a double energy position of the high count rate photons and also induces the tailing effect in the Gaussian peaks. Faster detectors have been developed for the elimination of the pile-up effect to a large extent. Also there are many pile-up rejection electronic techniques to remove such effects in the spectrum. Nevertheless, these techniques cannot reject the pile-up contribution fully because, the pile-up rate goes as $n^2\tau$, where n is true count rate and τ is the system dead time that cannot be reduced beyond a certain limit [4]. The electronics only rejects the counts which come within a certain interval and thus the detector does not represent absolute counts and therefore cannot be used for the measurement of absolute cross sections of such processes.

Now our method presents an alternative for eliminating the pile-up effect having the benefit of applicability for the measurement for absolute cross section measurements without using any expensive electronics. Unlike the electronic methods, this method works before the photons approach the detector and also has no limitation on the handling incident photon rate. It can differentiate between photons of different energies regardless of the interval between two incident photons. If the incident photon rate is high enough to generate the pile-up, the detector has difficulty in identifying the correct energy of photons when more than one photons fall on the detector at the same time (if the pile-up rejecter is not used). But if the pileup rejecter is used; once one photon falls on the detector then it leads to rejection of all other energy photons within its dead time, regardless of their energies and their numbers.

4.3. Getting back to the original intensity

As this method alters the original intensities of the peaks, it appears that this method is not useful for measuring the relative or absolute cross section measurements. But in fact, this method is equally good for the cross section measurements since the original intensity can be retrieved back experimentally by comparing the normalized spectra with and without an absorber. Once the thickness of the attenuator is correctly known, the original intensity of peaks will also be known too and hence this method is superior to the electronic methods.

4.4. Comparison with the electronic modules: effective detection for high energy transitions

As the attenuation coefficient for the higher energy transitions is much smaller than the transitions at low energy, therefore this technique can also be utilised for increasing the detection of higher energy photons by effectively lowering the count rate responsible for the low energy transitions. Hence, though the electronic techniques restrict count rate within 10^5/s for the X-ray detection with a solid state detector, we can keep the incident photon rate even much higher. This makes the absorber technique more efficient than the electronic pile-up rejection methods. For example, let the transitions are in the energy range 5–10 keV, where the intermediate heavy ions near $Z \sim 25$ can be studied. Suppose there are two peaks present at 5 and 10 keV having a total incident photon rate is 10^6 s^{-1} and a10 μm thick Al is selected as the absorber. For these two varieties of photons those at 10 keV are 1% of the total incident and rest are at 5 keV, hence, there are $99 \times 10^3 \rightarrow$ 5 keV and $10^3 \rightarrow 10$ keV photons. Therefore in one second for every 10^6 photons, 9×10^5 will not be detected as detector can detect only 10^6. While using an electronic module, out of 10^5 detected photons, 99×10^3 are 5 keV and 10^3 are 10 keV that realises 1% of total detected photons are 10 keV. Now, in the case of the absorber technique, we have selected Al (10 μm thick) that has absorption of ~40% for 5 keV and ~7% for 10 keV photons [13]. So there will be effectively about 98.46% of 5 keV photons and 1.54% of 10 keV photons of the total detected photons. Hence, 10 keV photons are detected more efficiently by a detector with the absorption technique in comparison to electronic modules. If the thickness of the absorber is increased this percentage will improve further for the high energy photon detection.

4.5. Life time measurements

This method can be used for the lifetime detection of the peaks after resolving them by the proposed technique. As the absorber always reduces a fixed ratio of incident ions, the transmitted intensity reaching the detector will be the same ratio throughout. Thus the decay curve characteristics will remain same. Hence, attenuating the intensity using an absorber does not affect the lifetime in the transmitted intensity data. Mathematically, Eq. (3) suggests that if a particular peak is measure at two different points n and m, their ratio remains fixed as follows:

$$\frac{I_n}{I_m} = \frac{I_{0n}}{I_{0m}} \quad (5)$$

Hence, the decay curves for lifetime measurements remains fully unaltered and so the shape of the decay. Fig. 3 shows a comparison of two simulated lifetime decay curves with and without absorber. Both the curves are fitted into a single exponential decay. Table 1 shows the fitting parameters of both the curves before and after the absorption of simulated data. Table 1 concludes that although the fitting parameter A and y_0 are different in the fitted curves because of absorption but the lifetime parameter τ remains the

Please cite this article in press as: G. Sharma et al., Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.062

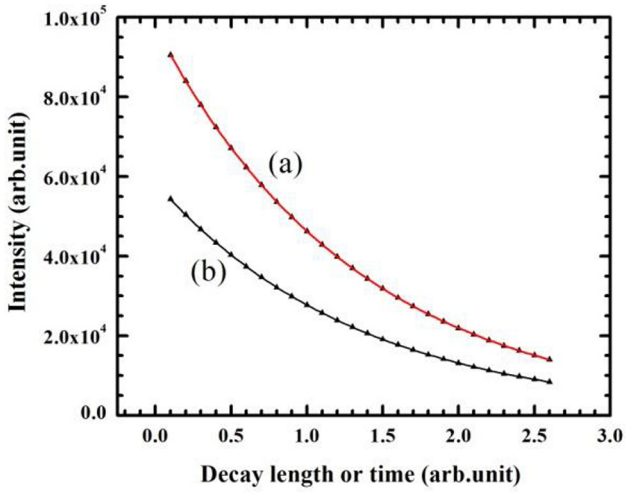


Fig. 3. Comparison of two life time curves (a) decay curve without any absorber, (b) decay curve after applying 10 μm Al absorber.

Table 1
Showing the fitting equation and the parameters of fitting for the curves, without and with Al absorber, fitted to the single exponential. Comparison of two curves shows the same decay-time (shown up to 5 decimal places) of both the curves. (τ is the lifetime of a state).

Equation; $y = A^*\exp(-x/\tau) + y_0$		Value	Standard error
Without absorber	y_0	−8.69E−1	3.90E−11
	A	9.76E4	3.05E−11
	τ	**1.33987**	1.29E−15
With absorber	y_0	−5.22E−1	2.50E−11
	A	5.86E4	1.96E−11
	τ	**1.33987**	1.38E−15

same. Thus, if this method is employed in an experiment it will represent the true lifetime of the transition. For the peaks which fit in other than exponential, say power law ($I = I_0\ t^{-a}$) fitting, intensity decay also shows the same behaviour as given in Eq. (5) and hence this method l holds good for them too.

5. Limitations and practicality of the technique

This method is only applicable if the intensity of the weaker peak is on the higher energy side of the spectrum and both the peaks are within the detector resolving power. If the weaker peak is on the low energy side then this will lead to the vanishing of the low energy peak and thus makes the method failed. However, the intense peak appears on the low energy side and weak peak is merged on the high energy side and hence the methods work well for the X-ray spectroscopy experiments. Therefore, many low intense peaks may exist in a spectrum those have been predicted theoretically but not yet been observed due to the merging of these peaks with the intense peaks. Hence, this technique will help experimentalists to classify those transitions in the spectrum. Hence the method aligns well with the experimental conditions.

Many a time the experimental conditions give rise to the X-ray continua [14] due to bremsstrahlung, recombination and 2E1 processes. They are broad continua spreading over a certain energy range and peaking at the centre. Hence, the present technique cannot be applied to study the physics of the X-ray continua. However, the peaks riding on the X-ray continuum can be studied, in particular, if the lines of interest are in the high energy half of the continuum.

6. Conclusion

The X-ray filtration technique is discussed in detail to resolve the low-intensity peaks veiled in their intense neighborhood peaks even though their line widths are expected to be resolved by the detector used; it also enables us to measure their lifetime, as the absorption does not alter the relative decay curves of the peaks. For removing the pile-up effect, the present method looks better when comparing with the presently available electronic techniques. Apart from the above mentioned benefits of the technique this method is a constructive technique and the original intensity can be retrieved from the measured data. This technique seems much better for detection of peaks in pile-up regions. However, apart from all the benefits, this technique can only be applied to resolve the merged peaks if the intense peak is on the lower energy side of the spectrum.

Acknowledgement

One of the authors (G.S.) thank for the University Grant Commission (UGC) for providing the valuable fellowship during his tenure as a doctoral student. We also acknowledge Prof. H.G. Berry for encouraging us and valuable discussions.

References

[1] M. Newville, Fundamentals of XAFS, in: Fundam, XAFS, 2004, pp. 1–2 (Chapter 1).
[2] R.K. Karn, C.N. Mishra, N. Ahmad, S.K. Saini, C.P. Safvan, T. Nandi, Note: Multi channel Doppler tuned spectrometer to study highly charged ions, Rev. Sci. Instrum. 85 (2014) 066110, http://dx.doi.org/10.1063/1.4884079.
[3] T. Nandi, P. Marketos, P. Joshi, R.P. Singh, C.P. Safvan, P. Verma, A. Mandal, A. Roy, R.K. Bhowmik, Lifetime of 1s2s2p 4P 5/2 in V20+ using beam-foil techniques, Phys. Rev. A 66 (2002) 052510, http://dx.doi.org/10.1103/PhysRevA.66.052510.
[4] Glenn F. Knoll, Ch-17, Linear and logic pulse functioning, 2009.
[5] J. Auerhammer, H. Genz, A. Kumar, A. Richter, Two-electron—one-photon transition in aluminum following double-K-shell ionization, Phys. Rev. A 38 (1988) 688–694.
[6] J.A. Tanis, J.M. Feagin, W.W. Jacobs, S.M. Shafroth, Comment on the comparison of observed two-electron-one-photon transition energies with calculated values, Phys. Rev. Lett. 38 (1977) 868–871.
[7] C. Stoller, W. Wolfli, G. Bonani, M. Stockli, M. Suter, Two-electron one-photon transitions into the doubly ionized K shell, Phys. Rev. A 15 (1977) 990.
[8] T. Aberg, K.A. Jamison, P. Richard, Origin of two-electron, one-photon K-X-ray transitions, Phys. Rev. Lett. 37 (1976) 63–65.
[9] T.K. Mukherjee, P.K. Mukherjee, Two electron one photon transitions in atomic systems, Z. Phys. D 42 (1997) 29–31.
[10] K.A. Sjiland, P. Kristiansson, Pile-up and defective pulse rejection by pulse shape discrimination in surface barrier detectors, Nucl. Instrum. Methods B 94 (1994) 333–337.
[11] R. Rolle, Comment on "the use of a pileup rejector in quantitative pulse spectrometry", Nucl. Instrum. Methods 114 (1973) 121.
[12] S.L. Blatt, J. Mahieux, D. Kohler, Elimination of pulsepileup ditortion in nuclear radiation spectra, Nucl. Instrum. Methods 60 (1967) 221–230.
[13] G. Weber, X-ray attenuation & absorption calculator, <http://web-docs.gsi.de/~stoe_exp/web_programs/x_ray_absorption/index.php>.
[14] N.J. Peacock et al., Quantitative spectroscopy of x-ray lines and continua in Tokamak, Rev. Sci. Instrum. 68 (1997) 1734.

Please cite this article in press as: G. Sharma et al., Theoretical approach for detection and lifetime measurement of obscured low cross-sectional processes by the X-ray absorption technique, Nucl. Instr. Meth. B (2017), http://dx.doi.org/10.1016/j.nimb.2017.05.062

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Theoretical investigation on the effect of nitrogen doping on the growth and field emission properties of the plasma-grown graphene sheet

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 J. Phys.: Conf. Ser. 836 012010

(http://iopscience.iop.org/1742-6596/836/1/012010)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 14.139.251.107
This content was downloaded on 06/06/2017 at 11:18

Please note that terms and conditions apply.

You may also be interested in:

A new approach to nanoporous graphene sheets via rapid microwave-induced plasma for energy applications
Taiwo Odedairo, Jun Ma, Yi Gu et al.

Study of Optical Properties of Zinc Oxide Thin Film Implanted with Nitrogen by Combinatorial Ion Implantation Techniques
Isao Sakaguchi, Yoshiyuki Sato, Hajime Haneda et al.

Theoretical Investigations of the Effect of N-Doping on Growth and Field Emission Properties of Carbon Nanotubes (CNTs)
I Santolia, A Tewari, S C Sharma et al.

Electron transport in graphene
S V Morozov, K S Novoselov and A K Geim

Theoretical Study of Epitaxial Graphene Growth on SiC(0001) Surfaces
Hiroyuki Kageshima, Hiroki Hibino, Masao Nagase et al.

Magnetron sputtering deposition of carbon nitride nanocolumns at low temperature
Bin Zhang, Yuanlie Yu and Junyan Zhang

Optimization of Immobilization of Nanodiamonds on Graphene
A Pille, S Lange, K Utt et al.

IOP Conf. Series: Journal of Physics: Conf. Series **836** (2017) 012010 doi:10.1088/1742-6596/836/1/012010

Theoretical investigation on the effect of nitrogen doping on the growth and field emission properties of the plasma-grown graphene sheet

N Gupta[a], S C Sharma, and R Gupta

Department of Applied Physics, Delhi Technological University, Delhi-110042, India

E-mail: [a)]nehagpt2105@gmail.com

Abstract. A theoretical model to study the impact of nitrogen-doping on the plasma-assisted growth and field emission properties of the graphene sheet (GS) has been developed. The model incorporates the charging rate of the GS, kinetics and energy balance of all plasma species i.e., electrons, positively charged ions and neutral atoms along with the nitrogen doping species, and growth rate of the GS. Numerical calculations on the effect of nitrogen doping on the thickness of the GS have been carried out for typical glow discharge parameters. It is found that the thickness of the GS decreases with nitrogen doping. The ramifications of nitrogen doping on the field enhancement factor of GS have likewise been examined on the premise of the above result. It is observed that the nitrogen doped GSexhibits better field emission as compared to undoped GS. Some of the results of the present investigation are in compliance with the experimental observations.

1. Introduction

Doping of graphene sheet (GS) using plasma enhanced chemical vapor deposition (PECVD) is considered as an effective method to deterministically control the growth of GS at low temperatures [1].Mehdipouret al.[2] and Sharma et al.[3] have developed the theoretical model to study the growth of carbon nano fibers and GS in the plasma environment respectively. Nitrogen is considered as one of the most popular candidate for doping of GS as it has similar radii as that of carbon and thus can replace carbon atoms in the graphene to induce n-type conduction in graphene [4]. Seoet al.[1] have reported that the thickness of the GS reduces upon addition of nitrogen to the $CH_4/Ar/H_2$ gas mixture. Moreover, the nitrogen doped GS exhibits high field enhancement factor [5]($\beta \approx \left(\lambda_{gr} / \tau_{gr}\right)$, where λ_{gr} is the height and τ_{gr} is the thickness of the GS) in comparison to undoped GS [6, 7].

2. Model

Following Sodha et al. [8], the present model considers the embryonic growth of GS in complex plasma consisting of electrons, positively charged ions and neutrals of CH_4 (denoted as A), neon (denoted as B), and nitrogen (denoted as C).Here, nitrogen is considered as the doping material.

The growth of GS in plasma environment can occur by many ways, namely, cluster formation, nucleation, coagulation, and growth of embryonic GS by condensation [9]. However, in the present investigation, we consider the embryonic growth of GS by condensation as we aim to explore the effect of nitrogen doping on the growth and field emission properties of the GS. Since the field

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

enhancement factor β of carbon nanostructures is less influenced by the height of the carbon nanostructures [10], we study the influence of nitrogen doping on the thickness of the GS.

2.1. Growth rate equation of the mass of graphene sheet

$$\alpha_{gr}\rho_{gr}\frac{d\tau_{gr}}{dt}=\sum_{k}^{A,C}M_k\chi_k I_{kgr}+\sum_{k}^{A,C}M_{ik}\chi_{ik}I_{ikgr}\,, \tag{1}$$

where $\alpha_{gr}\left(=l_{gr}\times\lambda_{gr}\right)$ is the area, l_{gr} is the length of the GS, and ρ_{gr} is the surface mass density of the GS. M_k and M_{ik} are the masses of atom and ion k (k refers to A and C type of atom or ion), respectively. Equation (1) describes the increase in thickness of the GS due to atomic $\left(I_{kgr}\right)$ and ion collection current $\left(I_{ikgr}\right)$ of species A and C. χ_k and χ_{ik} are the sticking coefficients of atom and ion of type k , respectively.

2.2. Kinetic equation of positively charged ions and neutral atoms

$$\frac{d\eta_{ik}}{dt}=\varepsilon_k\eta_k-\delta_k\eta_e\eta_{ik}-\eta_{gr}I_{ikgr}\,, \tag{2}$$

$$\frac{d\eta_k}{dt}=\delta_k\eta_e\eta_{ik}-\varepsilon_k\eta_k-\eta_{gr}\left(1-\chi_{ik}\right)I_{ikgr}-\eta_{gr}\chi_k I_{kgr}\,, \tag{3}$$

where η_k and η_{ik} are the number densities of atom and ion of type k (k refers to A, B or C type of atom or ion) , respectively.

Equation (2) represents the kinetics of positively charged ions, (here, k refers to A, B or C type of ion) in the plasma due to ionization of neutral atoms (ε_k is the coefficient of ionization), recombination of electrons and ions (δ_k is the coefficient of recombination), and ion collection current at the GS surface.

Equation (3) describes the kinetics of neutral atoms in the plasma due to electron-ion recombination, ionization of neutral atoms, ion collection current at the GS surface, and the accretion of neutral atoms A and C at the surface of GS. However, the last term of equation 3 i.e., the accretion of neutral atoms would not be accounted for the kinetics of neutral atom B.

3. Results and discussions

We have carried out the calculations to study the effect of nitrogen doping on the growth and field emission properties of the GS for typically glow discharge parameters. In particular, we have estimated the time variation of thickness of GS for undoped and nitrogen doped GS, and then premeditated the repercussion of nitrogen doping on the field enhancement factor β of the GS.

The figure 1 illustrates the variation of thickness of undoped and nitrogen doped GS with time. It can be seen from figure 1 that the thickness of GS decreases upon nitrogen doping. This can be attributed to the fact that the nitrogen atom, being pentavalent produces an extra electron on the GS surface which in turn increases the ionization of neutral atoms and thereby decreases the neutral atoms available for accretion. Since, the neutral atoms available for accretion decreases, the thickness of GS decreases on nitrogen doping. Moreover, the etching effect by nitrogen is additionally considered as the basis for decrease of thickness of nitrogen doped GS. The results of figure 1 are in consistence with the experimental observations of Seo et al. [1].

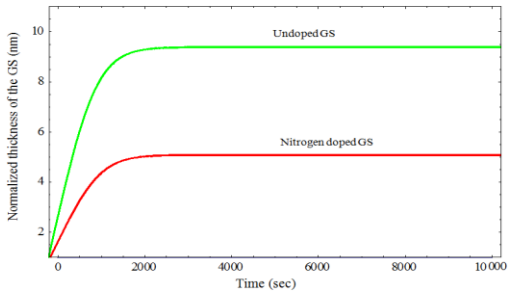


Figure 1. Variation of normalized thickness (in nm) of the undoped GS and nitrogen doped GS.

Using the above result, the field enhancement factor $\beta \approx \left(\lambda_{gr} / \tau_{gr}\right)$ is calculated using the saturated value of thickness corresponding to undoped and nitrogen doped GS. Since height $\left(\lambda_{gr}\right)$ has little influence on the field enhancement factor [10], it is assumed to be 1 μm for the calculation of β for undoped and nitrogen doped GS. It is found that the field enhancement factor is higher for nitrogen doped GS in comparison to undoped GS (cf. figure 2). The high field enhancement factor β of nitrogen doped GS is due to reduced thickness of GS upon nitrogen doping. The similar observation has been reported by Takeuchi et al. [6] and Palnitkar et al. [7].

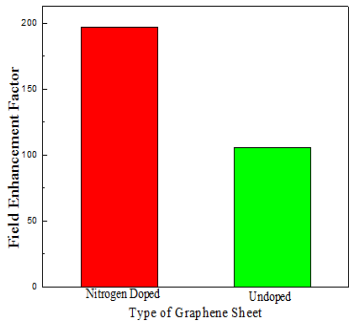


Figure 2. Field enhancement factor of undoped GS and nitrogen doped GS.

Acknowledgements
This work is partially supported by the Department of Science and Technology, Government of India.

References

[1] Seo D H, Kumar S, Rider A K, Han Z and Ostrikov K 2012 *Opt. Mat. Express* **2** 700
[2] Mehdipour H, Ostrikov K and Rider A E 2010 *Nanotechnology* **21** 455605
[3] Sharma S C and Gupta N 2015 *Phys. Plasmas* **22** 123517
[4] Rybin M, Pereaslavtsev A, Vasilieva T, Myasnikov V, Sokolov I, Pavlova A, Obraztsova E, Khomich A, Ralchenko V and Obraztsova E 2016 *Carbon* **96** 196
[5] Miller R, Lau Y Yand Booske J H 2007 *Appl. Phys. Lett.* **91** 074105
[6] Takeuchi W, Kondo H, Obayashi T, Hiramatsu M and Hori M 2011 *Appl. Phys. Lett.* **98** 123107
[7] Palnitkar U A, Kashid R V, More M A, Joag D S, Panchakarla L S and Rao C N R2010 *Appl. Phys. Lett.* **97** 063102
[8] Sodha M S, Misra S, Misra S K and Srivastava S 2010 *J. App. Phys.* **107** 103307
[9] Santolia I, Tewari A, Sharma S C and Sharma R 2014 *Phys. Plasmas* **21** 063508
[10] Xu Z, Bai X D and Wang E G 2006 *App. Phys. Lett.* **88** 133107

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Theoretical study to investigate the impact of plasma parameters on the catalyst nanoparticle growth

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2017 J. Phys.: Conf. Ser. 836 012024

(http://iopscience.iop.org/1742-6596/836/1/012024)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 14.139.251.107
This content was downloaded on 06/06/2017 at 11:19

Please note that terms and conditions apply.

You may also be interested in:

Control of growth mode of multiwalled carbon nanotubes
Nguyen Hong Quang and Do-Hyung Kim

Kinetic Calculations in Plasmas Used for Diamond Deposition
Pierre Bou, Jean Claude Boettner and Lionel Vandenbulcke

Carbon saturation of arrays of Ni catalyst nanoparticles of different size and patternuniformity on a silicon substrate
I Levchenko and K Ostrikov

Arc plasma synthesis of carbon nanostructures: where is the frontier?
Michael Keidar, Alexey Shashurin, Jian Li et al.

Low-and high-temperature controls in carbon nanofiber growth in reactive plasmas
H Mehdipour, K Ostrikov and A E Rider

Selective fabrication of quasi-parallel single-walled carbon nanotubes on siliconsubstrates
Xueshen Wang, Qunqing Li, Gang Zheng et al.

The use of microwave plasma-assisted CVD on nanostructured iron catalysts to growisolated bundles of carbon nanotubes
M B Assouar, M Dossot, S Rizk et al.

Minimizing the Gibbs–Thomson effect in the low-temperature plasma synthesis of thin Sinanowires
H Mehdipour, K Ostrikov, A E Rider et al.

A self-assembled synthesis of carbon nanotubes for interconnects
Zexiang Chen, Guichuan Cao, Zulun Lin et al.

IOP Conf. Series: Journal of Physics: Conf. Series **836** (2017) 012024 doi:10.1088/1742-6596/836/1/012024

Theoretical study to investigate the impact of plasma parameters on the catalyst nanoparticle growth

R Gupta[a], S C Sharma and N Gupta

Department of Applied Physics, Delhi Technological University, Delhi-110042, India

Email: [a)]ravigpt0709@gmail.com

Abstract. The plasma kinetics based model is adopted to elucidate the effect of plasma parameters on the nucleation and growth mechanism of catalyst nanoparticle. The present model considers the plasma processing of thin catalyst film, power equalization at the film surface, flux and kinetics of plasma species (electrons, ions, and neutrals). In our investigation, it is found that catalyst nanoparticle diameter decreases with increase in ion number density in plasma. Moreover, it is also found that catalyst film thickness significantly affect the catalyst nanoparticle size i.e., catalyst nanoparticle diameter increases with catalyst film thickness. In addition, it is observed that the substrate temperature increases during the plasma processing and finally achieve saturation. Our theoretical results are in good agreement with the experimental results.

1. Introduction

The catalyst nanoparticles are the fundamental site for the nucleation of nanostructures and significantly influence the structure, shape and size of the nanostructure [1]. These catalyst nanoparticles can be synthesized using the plasma technique, i.e., plasma sputtering and etching of the thin film[2]or by thermal annealing of the thin film deposited on substrate surface. During the plasmatreatment of thin film, plasma parameters extensively affect the catalyst nanoparticle span (diameter). Chang et al. [3] experimentally observed that the catalyst nanoparticle size obtaining from thin catalyst film decreases with increases in plasma power.

So far only thermodynamics based theoretical models are available to depict the size of the catalyst nanoparticle obtained from the thin catalyst film. In the present paper, in section 2, we have devised the plasma based model to elucidate the growth mechanism of catalyst nanoparticle resulting from the etching and sputtering of the thin film in the reactive plasma, i.e. plasma treatment of the thin catalyst film. The results of the model are discussed in the section 3 and finally conclusion is given in section 4.

2. Model

The model acknowledges the kinetics and energy fluxes of various plasma species (electrons, positively charged ions and neutrals), power equalization at the catalyst film-substrate surface, and heat radiations. The present model accounts the reactive plasma consisting electrons, ions and neutrals of hydrogen. It is assumed that, the spherical catalyst nanoparticles are formed as a result of sputtering and etching of the catalyst film during the plasma treatment.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

2.1. Kinetic equation of electrons, positively charged ions, and neutrals in the reactive plasma

$$\dot{n}_e = \beta_1 n_1 - \alpha_1 n_e n_{i1} - \gamma_e n_a j_{ecf} \tag{1}$$

$$\dot{n}_{i1} = \beta_1 n_1 - \alpha_1 n_e n_{i1} - n_a j_{i1cf} - J_{ad} + J_{desp} \tag{2}$$

$$\dot{n}_1 = \alpha_2 n_e n_{i1} - \beta_1 n_1 + n_a (1-\gamma_{i1}) j_{i1} - n_a \gamma_1 j_1 \tag{3}$$

where subscript 1 denotes the hydrogen, β_1 is the coefficient of ionization of the constituent neutral atoms due to external field, n_2 is the number density of neutral atom, n_e is the electron number density, n_{i1} is the ion number density, $\alpha_j(T_e) = \alpha_{j0}(300/T_e)^k \, cm^3/\sec$ is the coefficient of recombination of electrons and positively charged ions, $k = -1.2$ is a constant, η_a is the number of catalyst particle per unit area γ_{i1} is the ion sticking coefficient, γ_2 is the sticking coefficient of neutral atoms J_{ad} is the adsorption flux onto the catalyst substrate surface, J_{desp} is the desorption flux from the catalyst-substrate surface, j_{ecf}, $j_{i\alpha cf}$ and $j_{\alpha f}$ are the collection currents on the film surface due to electrons, ions and neutrals, respectively [4].

Equations (1), (2), and (3) indicate the balance of electrons, ions, and neutral atoms number density in the reactive plasma. The first of equation (1)denotes the gain in electron density per unit time due to ionization of neutral atoms; second and third terms refers to the decay rate due to electron-ion recombination and electron collection current at the film surface, respectively.

The first term of equation (2) describes the gain in ion number density due to ionization of neutral atoms. The second and the third terms represent the decay rate of positively charged ion density due to electron-ion recombination and ion collection current at the surface of the film, respectively. The fourth and fifth terms indicate adsorption and desorption of ions to/from the film surface, respectively

The First and third terms of equation (3) show the growth rate of neutral density due to electron-ion recombination and neutralization of ions collected at the film surface, respectively. The second and fourth terms represents the decay rate of neutral density due to ionization and accumulation of neutrals on the film surface, respectively.

2.2. The power equalization at the catalyst film surface

$$\dot{H}_{input} = \dot{H}_g + \dot{H}_{output} \tag{4}$$

where $\dot{H}_{in}$ is the input energy per unit time towards the catalyst film, $\dot{H}_g$ is the energy gained per unit time during the plasma treatment and $\dot{H}_{out}$ is the output energy or energy loss per unit time [5]. Various plasma processes contribute to these energies.

$$\dot{H}_{input} = \left[j_{ecf} \xi_e^{cf} + j_{i1cf} \xi_{i1}^{cf} + j_{1cf} \xi_1^{cf} \right] + A_{cf} \sigma \left((\vartheta_{srd})^4 \in_{srd} - (\vartheta_S)^4 \in_{cf} \right) + \frac{3}{2} k_B \vartheta_S (1-\kappa_{i1}) j_{i1cf} \tag{5}$$

$$\dot{H}_g = \frac{d}{dt}\left(m_{cf} S_f \vartheta_S \right) = A_{cf} S_{cf} \rho_{cf} \left[\tau_{cf} \frac{d\vartheta_s}{dt} + \eta_a \vartheta_s \frac{d}{dt}\left(\frac{1}{6} \pi d^3 \right) \right] \tag{6}$$

$$\dot{H}_{output} = \Lambda_{ef} (1-\kappa_e) \left(\xi_{ecf}^S - \frac{3}{2} k_B \vartheta_S \right) j_{ecf} + \sigma A_{cf} \left(\vartheta_s^4 \in_{cf} - \vartheta_{srn}^4 \in_{srn} \right) + \left[\upsilon_0 \sigma \varphi_1 E_B + \varphi_{i1} \left[y_d(E_i) \right] (1-\theta) \right] A_{cf}, \tag{7}$$

where A_{cf} is the catalyst film surface area, τ_{cf} is the catalyst film thickness, d is the catalyst nanoparticle diameter (it is assumed that catalyst nanoparticle of spherical morphology are formed as result of plasma treatment of thin catalyst film), ϑ_s is the substrate temperature (in the present model temperature of substrate and thin catalyst film is assumed to be same), ϑ_{cf} is the mass of catalyst film,

κ_{i1} and κ_1 are the ion sticking coefficient and neutral atom sticking coefficient, respectively, s_f is the specific heat of the film, $\in_{srn}$ is the emissivity of the surrounding, $\in_{cf}$ is the emissivity of the film, ξ_e^{cf} , ξ_{i1}^{cf} , and ξ_1^{cf} are the mean energy of electrons, ions and neutrals, respectively collected by film [6], Λ_{ef} is the energy loss during collision with film, υ_0 is the no. of sites available, σ_{ads} is the cross-section area, φ_a and φ_{ia} are neutral and ion fluxes, respectively, E_B is the binding energy of the material, σ is the Stefan's constant and θ is the total surface coverage.

The first term in equation (5) represents the power given to the film surface by the electrons, ions and neutral atoms, the second term denotes power towards the film surface due to heat radiations from surrounding, and the third term is power collected by film surface due to formation of neutrals at the film surface. The equation (6) describes the power gained during the plasma treatment of thin catalyst film. The input plasma power causes the thin catalyst film into nanoparticles and also causes the increase in substrate temperature. The equation (7) describes the power loss due to sticking and elastic collision of electrons with catalyst film surface, power output from the surface to the surrounding in the form of radiations, power loss in etching and sputtering of thin film in the presence of plasma, respectively.

3. Results and discussion

The model acknowledges the different physical process responsible for the formation of catalyst nanoparticle from thin catalyst film in the presence of plasma and hence the model relates the diameter of the catalyst nanoparticle formed with the plasma parameters such that number density of the electrons, ions, and neutrals, plasma power, and thin film thickness.

The figure1 shows the time evolution of catalyst nanoparticle diameter for various thin film thickness. From figure 1, it can be seen that diameter of the catalyst nanoparticle increases with thickness of the thin film.

The figure 2 shows variation of catalyst nanoparticle diameter with plasma power. As plasma power increases the catalyst nanoparticle diameter decreases. This is attributed to the fact that with increase in plasma power highly energetic plasma species (electrons, ions, andneutrals) are created in the reactive plasma that leads to higher etching and sputtering of the catalyst thin film resulting in the lesser catalyst nanoparticle diameter.

The figure 3 shows the variation of substrate temperature during the plasma treatment of the thin catalyst film. The substrate temperature increases during the plasma treatment and achieves the saturation after sometime. The highly energetic plasma species transfer their energy to the film surface during the bombardment and collisional processes with the film surface and thus increase the film surface/substrate temperature. These theoretical findings are in good agreement of the experimental observations of Poa et al. [7], Chang et al. [3], and Srivastava et al. [8].

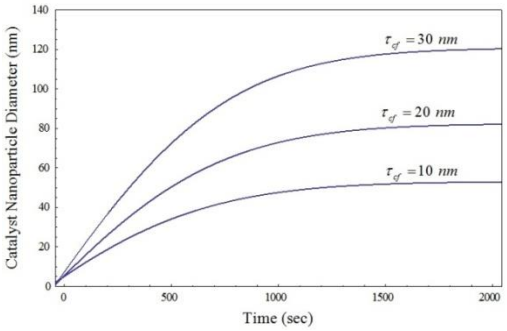


Figure 1. The time evolution of catalyst nano-particle diameter for various catalyst film thicknesses.

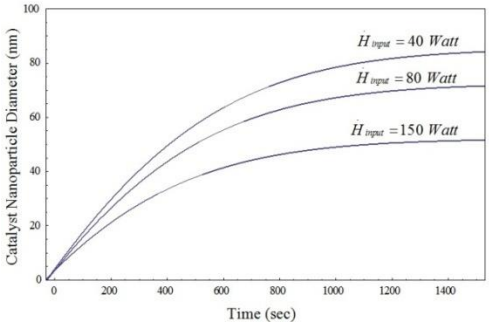


Figure 2. The time evolution of catalyst nanoparticle for different input plasmapower.

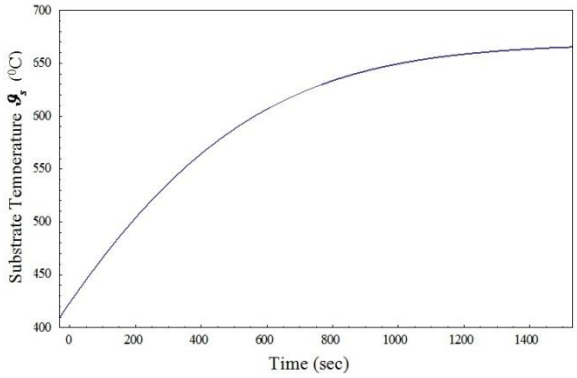


Figure 3. The time variation of substrate temperature during the plasma processing (plasma treatment) of the thin catalyst film in the plasma containing electrons and ions and neutrals of hydrogen.

4. Conclusions

An analytical model to relate the catalyst nanoparticle diameter with plasma parameters (number density of plasma species, their temperatures, and plasma power) and catalyst film thickness during the plasma treatment catalyst film has been developed. It is found that diameter of spherical catalyst nanoparticles increases with catalyst thin film thickness and decreases with plasma power. Moreover, we found that substrate temperature also increases during the plasma treatment of thin film and after some time achieves the saturation. The present work could foresee the catalytic synthesis of the carbon nanostructure in near future.

References

[1] Yang R T and Chen J P 1989 *J. Catal* **115** 52
[2] Yudasaka M, Kikuchi R, Matsui T, Ohki Y, Yoshimura S and Ota E 1995 *Appl. Phys. Lett.* **67** 2477
[3] Chang S C, Lin T C and Pai C Y 2007 *Microelectron. J.* **38** 657
[4] Shrama S C and Gupta N 2015 *Phys. Plasmas* **22** 123517
[5] Kersten H, Deustch H, Steffen H, Kroesen G M W and Hippler R 2001 *Vacuum* **63** 385
[6] Sodha M S and Guha S 1971 *Physics of Colloidal Plasmas* (New York: John Wiley & Sons) p 219
[7] Poa C H P, Henley S J, Chen G Y, Adikaari A A D T, Giusca C E and Silva S R P 2005 *J. Appl. Phys.* **97** 114308
[8] Srivastava S K, Shukla A K, Vankar V D and Kumar V 2005 *Thin Solid Films* **492** 124

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

Thermal modelling of three stage vapour compression cascade refrigeration system using entropy generation principle for reducing global warming and ozone depletion using ecofriendly refrigerants for semen preservation

R S Mishra

Department of Mechanical, Production & Industrial and Automobiles Engineering
Delhi Technological University Delhi
E mail: rsmishra@dce.ac.in , rsmishra@dtu.ac.in

Abstract

This paper mainly deals with thermodynamic analysis of three stages cascade vapour compression refrigeration systems using ecofriendly refrigerants used for low temperature applications. The effect of thermal performance parameters (i.e. approaches, condenser temperature, and temperature variations in the evaporators) on the first law thermal performances COP_{System} and also in terms of second law efficiency of the cascade system (exergetic efficiency) and System exergy destruction ratio (SEDR) have been optimized thermodynamically using entropy generation principle.
The utility of R1234ze and R1234yf and in the high temperature circuits and new ecofriendly refrigerants in the intermediates circuits and R134a or R404a in the low temperature cascade circuit have been optimized. It was observed that in the low temperature (between- 50oC to -100oC) applications. It was observed that the best combination in terms of R1234ze-R134a-R404a gives better thermal performance than using R1234yf-R134a-R404a**.** Similarly other combination in terms of R1234ze-R134a-R404a gives better thermal performance than using R1234ze-R1234yf-R404a.

© 2017 ijrei.com. All rights reserved

Keywords: Three stages Cascade VCR, Energy-Exergy Analysis, and Thermodynamic Optimization

1. Introduction

Low-temperature refrigeration systems are typically required for low temperature range from –40°C to –100°C for applications chemical, food, pharmaceutical, and other industries. For low temperature applications, single-stage/Double stages systems with reciprocating compressors are not feasible due to higher pressure ratios. High pressure ratio means high discharge and oil temperatures with low volumetric Compressor efficiencies which results low first law efficiency in terms of COP values. Vapor compression cascade refrigeration systems be used to achieve low temperatures applications in such a manner , where series of single-stage units are used that are thermally coupled through evaporator/condenser cascading. Each circuit has a different ecofriendly refrigerant suitable for that temperature. The high-temperature circuit uses high boiling point refrigerants such as R-1234ze, and R1234yf, R 717 and R152a and Intermediate circuit's ecofriendly refrigerants such as R-1234yf, R134a, R407c R125, etc. are used. Similarly for the low-temperature cascade circuit low boiling refrigerants are used.
In a vapour compression cascade refrigeration system where low pressure vapour in the evaporator stage is compressed and then recycled for condensation in the evaporator of a previous stage, the improvement of passing the low pressure refrigerant's vapour from an evaporator stage through a heat exchanger to heated vapour to ambient temperature, compressing said heated vapour, removing the compressor work by passing said compressed vapour through a evaporator, then cooling is by compressed

Corresponding Author: R.S. Mishra
Email Id: rsmishra@dce.ac.in

vapour by passing it through heat exchanger in heat exchanger at a low pressure vapour, and condensing it in the evaporator of the next higher temperature cycle of the cascade system and recycling the liquid to the low pressure evaporator stage. Most of the investigators even did not found the effect of various approaches on the system first and second law performances on three stages cascade refrigeration systems and effect of ecofriendly refrigerants in the intermediate temperature circuit.

This paper mainly deals with thermal performances in terms of exergy destruction ratio by thermodynamically analyzing the use of R1234yf and R1234ze in the high temperature circuits and R134a/ R404a eco-friendly refrigerants in the low temperature circuit.

1.1 Literature Review

B. Agnew, et.al [1] considered finite time analysis of a cascade refrigeration system using alternative refrigerants. Bansal P.K [2] used thermodynamic analysis of carbon dioxide–ammonia (R744–R717) cascade refrigeration system. Bhattacharyya et al [3] carried out thermal modelling of a cascade refrigeration–heat pump system by incorporating both internal and external irreversibility's. Cabell et al [4] developed steady-state modelling of a single stage vapours compression plant. James M. Calm, [5] considered the next generation of refrigerants in VCR. Ecir et al [6] used thermal modelling techniques for predicting thermo-physical properties of refrigerants. Kilicarslan [7] carried out experimental investigation of a different type of two-stage vapour compression cascade refrigeration system. Jain et al [8] showed that carbon dioxide is a potential low temperature refrigerant.. Pearson [9] traced the development of the old carbon dioxide systems Samant [10] design and development of two stage cascade refrigeration system using CO2 as LTC refrigerant and Propane as HTC refrigerant. Resat S. et al [11] exergy based thermo-economic optimization of subcooled and superheated vapour compression refrigeration cycle. Wang et al [12] potential benefits of compressor cooling. Tzong-Shinge et al [13] comprehensive investigation of numerical methods in simulating a steady-state vapor compression system. Mukhopadhyay, [14] used CO2–C3H8 cascade refrigeration–heat pump system: and carried out numerical verification of heat exchanger inventory optimization Mishra [15–16] examined the performance of a Four stage cascade refrigeration systems using hydrocarbon as refrigerants and evaluated thermal performance of four stages cascade vapour Compression System and theoretical study of a different type of two-stage vapour compression refrigeration system.

The above investigators did not studies the thermodynamic performances in terms of COP and exergetic efficiency and system exergy destruction ratio (SEDR) for very low temperature application using ecofriendly refrigerants in the low temperature circuits and R1234ze in higher temperature circuit, and R1234yf in intermediate temperature circuit and R134a in low temperature circuit.

2. Description of Three Stage Cascade Systems

The following assumptions have been taken for analyzing three stages cascade vapour compression system for low temperature applications. The cooling load is considered to be 70 [kW‖ Temperature of condenser is to be 50^0C and temperature of low circuit evaporator to be -100^0C, The temperature of high temperature evaporator to be -10^0C, temperature of secondary intermediate cascade evaporator is to be -50^0C, The effect of temperature overlapping (approach means temperature difference between cascade condenser and cascade evaporator is 10^0C in each stage is considered.

3. Results and Discussions

For Cascade vapor compression refrigeration System using ecofriendly refrigerants R-1234ze-R1234yf-R134a in high temperature circuit, intermediate temperature circuit and low temperature circuit, for reducing global warming and ozone depletion following Input values are taken for numerical computation of model developed by author

Compressor efficiency in each temperature circuit is 80%. Condenser temperature in High temperature circuit is 50^0C, temperature of High temperature evaporator is 0^0C. Temperature of intermediate temperature evaporator is -50^0C and Temperature of low temperature evaporator is -100^0.

Approach (i.e. Temperature Overlapping) means the temperature difference between cascade condenser (intermediate temperature circuit) minus high temperature evaporator circuit and another approach (i.e. temperature overlapping between low temperature cascade condenser temperature minus evaporator temperature in intermediate temperature circuit is ranging from 0^0C to 10^0C.

Table-1(a) to 1(b) showing the first law efficiency of the three stage cascade refrigeration system usingR1234ze in the high temperature circuit and R1234yf in the secondary cascade intermediate temperature circuit and R134a in a lower temperature circuit. It was observed that the use of R1234ze in higher temperature circuit gives better thermodynamic performance than using R1234yf in high temperature circuit up to temperature range of -100^0C.

However for low temperature applications The best combination of three stage cascade is system is to be R1234ze-R134a-R404a .The other refrigerants such as R123 and hydrocarbons containing chlorine and R152, R600 and R290 are also flammable in nature can be considered by adopting safety measures while R717 in toxic in nature. Therefore R1234ze-R1234yf-R134a is most suitable for practical applications.

Table-1(a) Variation of system thermal performance of three stages cascade vapour compression refrigeration systems with ecofriendly refrigerant in high temperature circuit

Ecofriendly Refrigerant in Low Temperature Circuit	System First Law Efficiency (COP)	Exergy Destruction Ratio EDRSystem	Exergetic Efficiency/ Second Law efficiency/ EFFSecond
R1234ze	0.5074	1.728	0.3666
R1234yf	0.4963	1.789	0.3586
R717	0.5249	1.637	0.3792
R152a	0.5201	1.661	0.3758

Table-1(b).Variation of system thermal performance along with circuit thermal performance of three stages cascade vapour compression refrigeration systems with ecofriendly refrigerant in high temperature circuit

Ecofriendly Refrigerant in Low Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPITC	COPLTC
R1234ze	0.5074	3.215	2.204	1.79
R1234yf	0.4963	3.027	2.204	1.79
R717	0.5249	3.546	2.204	1.79
R152a	0.5201	3.451	2.204	1.79

Table-2(a).Variation of thermal performance of three stages cascade vapour compression refrigeration systems with ecofriendly refrigerant in intermediate temperature circuit

Ecofriendly Refrigerant in Intermediate Temperature Circuit	System First Law Efficiency (COP)	Exergy Destruction Ratio EDRSystem	Efficiency Second
R1234yf	0.5074	1.728	0.3666
R134a	0.5169	1.677	0.3735
R407c	0.4795	1.887	0.3464
R410a	0.5152	1.682	0.3723
R404a	0.4966	1.787	0.3588

Table-2(b).Variation of thermal performance of three stages cascade vapour compression refrigeration systems with ecofriendly refrigerant in intermediate temperature circuit

Ecofriendly Refrigerant in Intermediate Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPITC	COPLTC
R1234yf	0.5074	3.215	1.79	2.204
R134a	0.5169	3.215	1.79	2.294
R407c	0.4795	3.215	1.79	1.961
R410a	0.5152	3.215	1.79	2.278
R404a	0.4966	3.215	1.79	2.107

Table-3(a).Thermal performance of three stages cascade vapor compression refrigeration systems with ecofriendly refrigerant in low temperature circuit

Ecofriendly Refrigerant in Low Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
R134a	0.5074	1.728	0.3666
R404a	0.4971	1.784	0.3592
R600a	0.5123	1.701	0.3702
R290	0.510	1.714	0.3685
R600	0.5148	1.688	0.3720
Ethylene	0.4781	1.895	0.3454
Ethane	0.4951	1.795	0.3578

Table-3(b).Variation of thermal performance of three stages cascade vapour compression refrigeration systems with ecofriendly refrigerant in low temperature circuit

Ecofriendly Refrigerant in Low Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
R134a	0.5074	3.215	2.204	1.79
R404a	0.4971	3.215	2.204	1.724
R600a	0.5123	3.215	2.204	1.822
R290	0.510	3.215	2.204	1.807
R600	0.5148	3.215	2.204	1.839
Ethylene	0.4781	3.215	2.204	1.607
Ethane	0.4951	3.215	2.204	1.711

Table-4(a) Variation of thermal performance of three stages cascade vapor compression refrigeration systems

Effect of Approach in Low Temperature Circuit	System First Law Efficiency (COP)	EDR System	Efficiency Second
0	0.570	1.428	0.4119
2.5	0.5535	1.50	0.3999
5.0	0.5376	1.574	0.3884
7.5	0.5222	1.650	0.3773
10.0	0.5074	1.728	0.3666

Table-4(b) Variation of thermal performance of three stages cascade vapor compression refrigeration systems

Effect of Approach in Low Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
0	0.570	3.215	2.204	2.247
2.5	0.5535	3.215	2.204	2.116
5.0	0.5376	3.215	2.204	1.998
7.5	0.5222	3.215	2.204	1.889
10.0	0.5074	3.215	2.204	1.79

Table-5(a) Variation of thermal performance of three stages cascade vapour compression refrigeration systems

Effect of Approach in Low Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
0	0.5675	1.439	0.410
2.5	0.5522	1.507	0.399
5.0	0.5371	1.577	0.388
7.5	0.5221	1.651	0.3773
10.0	0.5074	1.728	0.3666

Table-5(b) Variation of thermal performance of three stages cascade vapour compression refrigeration systems

Effect of Approach in Low Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPITC	COPLTC
0	0.5675	3.215	2.844	1.79
2.5	0.5522	1.507	2.633	1.79
5.0	0.5371	1.577	2.497	1.79
7.5	0.5221	1.651	2.345	1.79
10.0	0.5074	1.728	2.204	1.79

Table-6(a) Variation of condenser temperature on thermal performance of three stages cascade vapour compression refrigeration systems

Condenser temperature in High Temperature (R1234ze) Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
60	0.4530	2.055	0.3273
55	0.4804	1.881	0.3471
50	0.5074	1.728	0.3666
45	0.5340	1.592	0.3859
40	0.5606	1.469	0.4051
35	0.5873	1.356	0.4244
30	0.6143	1.253	0.4438
25	0.6416	1.157	0.4636

Table-6(b) Variation of condenser temperature on thermal performance of three stages cascade vapour compression refrigeration systems

Condenser temperature in High Temperature (R1234ze) Circuit	System First Law Efficiency (COP)	COPHTC	COPITC	COPLTC
60	0.4530	2.407	2.204	1..79
55	0.4804	2.779	2.204	1..79
50	0.5074	3.215	2.204	1..79
45	0.5340	3.737	2.204	1..79
40	0.5606	4.379	2.204	1..79
35	0.5873	5.192	2.204	1..79
30	0.6143	6.264	2.204	1..79
25	0.6416	7.750	2.204	1..79

Table-7(a) Variation of evaporator temperature on thermal performance of three stages cascade vapor compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
20	0.4698	1.946	0.3395
15	0.4847	1.855	0.3502
10	0.4958	1.791	0.3583
5	0.5033	1.75	0.3637
0	0.5074	1.728	0.3666
-5	0.5081	1.724	0.3671
-10	0.5057	1.737	0.3654
-15	0.5003	1.766	0.3615
-20	0.4923	1.811	0.3557

Table-7(b) Variation of evaporator temperature on thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
20	0.4698	6.446	1.356	1.79
15	0.4847	5.287	1.533	1.79
10	0.4958	4.421	1.729	1.79
5	0.5033	3.75	1.951	1.79
0	0.5074	3.225	2.204	1.79
-5	0.5081	2.78	2.497	1.79
-10	0.5057	2.42	2.844	1.79
-15	0.5003	2.117	3.261	1.79
-20	0.4923	1.86	3.777	1.79

Table-8(a) Variation of evaporator temperature on thermal performance of three stages cascade vapor compression refrigeration systems

S.No	Evaporator temperature in Low Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
1	-80	0.6759	1.72	0.3677
2	-85	0.6312	1.708	0.3693
3	-90	0.5882	1.705	0.3697
4	-95	0.5470	1.712	0.3687
5	-100	0.5074	1.728	0.3666

Table-8(b) Variation of evaporator temperature on thermal performance of three stages cascade vapor compression refrigeration systems

Evaporator temperature in Low Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
-80	0.6759	3.215	2.204	3.323
-85	0.6312	3.215	2.204	2.810
-90	0.5882	3.215	2.204	2.40
-95	0.5470	3.215	2.204	2.067
-100	0.5074	3.215	2.204	1.79

Table-9(a) Variation of condenser temperature on thermal performance of three stages cascade vapor compression refrigeration systems

Condenser temperature in High Temperature (R1234ze) Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
60	0.4530	2.055	0.3273
55	0.4804	1.881	0.3471
50	0.5074	1.728	0.3666
45	0.5340	1.592	0.3859
40	0.5606	1.469	0.4051
35	0.5873	1.356	0.4244
30	0.6143	1.253	0.4438
25	0.6416	1.157	0.4636

Table-9(b) Variation of evaporator temperature on thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
-50	0.5074	3.215	2.204	1.79
-45	0.5092	3.215	2.529	1.613
-40	0.5090	3.215	2.921	1.461
-35	0.5067	3.215	3.404	1.328
-30	0.5023	3.215	4.011	1.211

Table-10(a) Variation of condenser temperature on thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
20	0.4698	1.946	0.3395
15	0.4847	1.855	0.3502
10	0.4958	1.791	0.3583
5	0.5033	1.75	0.3637
0	0.5074	1.728	0.3666
-5	0.5081	1.724	0.3671
-10	0.5057	1.737	0.3654
-15	0.5003	1.766	0.3615
-20	0.4923	1.811	0.3557

Table-10(b) Variation of condenser temperature on thermal performance of three stages cascade vapour compression refrigeration systems

Condenser temperature in High Temperature (R1234ze) Circuit	System First Law Efficiency (COP)	COPHTC	COPITC	COPLTC
60	0.4530	2.407	2.204	1..79
55	0.4804	2.779	2.204	1..79
50	0.5074	3.215	2.204	1..79
45	0.5340	3.737	2.204	1..79
40	0.5606	4.379	2.204	1..79
35	0.5873	5.192	2.204	1..79
30	0.6143	6.264	2.204	1..79
25	0.6416	7.750	2.204	1..79

Table-11(a) Variation of condenser temperature on thermal performance of three stages cascade vapor compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
20	0.4698	1.946	0.3395
15	0.4847	1.855	0.3502
10	0.4958	1.791	0.3583
5	0.5033	1.75	0.3637
0	0.5074	1.728	0.3666
-5	0.5081	1.724	0.3671
-10	0.5057	1.737	0.3654
-15	0.5003	1.766	0.3615
-20	0.4923	1.811	0.3557

Table-11(b) Thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
20	0.4698	6.446	1.356	1.79
15	0.4847	5.287	1.533	1.79
10	0.4958	4.421	1.729	1.79
5	0.5033	3.75	1.951	1.79
0	0.5074	3.225	2.204	1.79
-5	0.5081	2.78	2.497	1.79
-10	0.5057	2.42	2.844	1.79
-15	0.5003	2.117	3.261	1.79
-20	0.4923	1.86	3.777	1.79

Table- 12 (a) Thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator Temperature nt in Low Temperature Circuit	System First Law Efficiency (COP)	EDR System	EFF Second
-80	0.6759	1.72	0.3677
-85	0.6312	1.708	0.3693
-90	0.5882	1.705	0.3697
-95	0.5470	1.712	0.3687
-100	0.5074	1.728	0.3666

Table- 12 (b) Thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator Temperature in Low Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
-80	0.6759	3.215	2.204	3.323
-85	0.6312	3.215	2.204	2.810
-90	0.5882	3.215	2.204	2.40
-95	0.5470	3.215	2.204	2.067
-100	0.5074	3.215	2.204	1.79

Table- 13 (a) Thermal performance of three stages cascade vapour compression refrigeration systems

Evaporator temperature in intermediate Temperature Circuit	System First Law Efficiency (COP)	COPHTC	COPLTC	COPLTC
-50	0.5074	3.215	2.204	1.79
-45	0.5092	3.215	2.529	1.613
-40	0.5090	3.215	2.921	1.461
-35	0.5067	3.215	3.404	1.328
-30	0.5023	3.215	4.011	1.211

From Tables-3: it was observed that as condenser temperature decreases with increasing First law efficiency of cascade system in terms of overall system COP. Similarly the System EDR is decreases and also second law efficiency/exergetic efficiency increases with decreasing condenser temperature. Regarding First law efficiency in terms of circuit COP of hot fluid circuit is increases. There will not be effect on other circuit first law efficiencies (COP). The exergetic efficiency and Overall COP of system using R1234ze is higher than using R1234yf in the high temperature circuit while system EDR is increases. As decreasing High temperature evaporator temperature, the first law efficiency (i.e. overall COP of system) and second law efficiency (i.e. exergetic efficiency of whole system) is increasing up to maximum value at the evaporator temperature of 25oC and then decreases rapidly.

As decreasing High temperature evaporator temperature the system EDR first decreasing up to decreasing evaporator temperature and then further constant and then increasing and optimum becomes at evaporator temperature at 25oC in both cases using R1234yf and R1234ze in the high temperature circuit. However first law efficiency (COP) of hot fluid circuit is decreases and COP of primary intermediate fluid circuit is increases.. As evaporator temperature decreases, the first law efficiency (overall COP) and second law efficiency (exergetic efficiency) of the system is increases and maximum efficiency is obtained at evaporator temperature of -10 oC and also cop of primary intermediate temperature circuit is decreases and secondary intermediate temperature circuit COP is increases.

As increasing evaporator temperature overall cop is increases while System EDR first decreases while second law efficiency is increases. The maximum efficiency was obtained at -50 oC and then decreases as increasing evaporator temperature. Secondary intermediate circuit COP is increases while low temperature circuit COP is decreases. Table-6(b): Performance variation with cascade evaporator temperature in the secondary intermediate temperature circuit temperature of four stage cascade vapour compression refrigeration system using R1234yf in High temperature circuit and R404a in the low temperature circuit Table-7(a) to 7(d) are showing the effect of low temperature evaporator on the system first and second law performances and system exergy destruction ratio and various circuit first law performances. As low temperature evaporator temperature is decreasing, the first law and second law efficiencies are increasing and exergy destruction ratio is decreasing and high temperature circuit first law performances in increasing.

4. Conclusion

The following conclusions were drawn while analysing three stages cascade refrigeration systems.

(i) The use of R1234ze has Global warming potential (GWP) = 6 gives better thermal performance than R1234yf of 4GWP in the higher temperature circuit of three stages cascade refrigeration system.
(ii) There is significant performance improvements using R134a as compared with R1234yf in the intermediate temperature circuit with R134a in lower temperature circuit.
(iii) The second law efficiency using R134a in the low temperature evaporator circuit gives better performance as compared with R410a in intermediate temperature circuit.

5. References

[1] B. Agnew, et.a.,[2004], A finite time analysis of a cascade refrigeration system using alternative refrigerants, Applied Thermal Engineering, 24, , 2557–2565.

[2] P. K. Bansal,[2008], Thermodynamic analysis of an R744–R717 cascade refrigeration system, international journal of refrigeration 31, , 45-54.

[3] Souvik Bhattacharyya, [2007], Exergy maximization of cascade refrigeration cycles and its numerical verification for a trans critical CO2-C3H8 system, International Journal of Refrigeration, 30, , 624-632.

[4] R. Cabello,[2005], Simplified steady-state modelling of a single stage vapour compression plant. Model development and validation, Applied Thermal Engineering, 25, 1740–1752.

[5] James M. Calm, [2008],The next generation of refrigerants – Historical review, considerations, and outlook, International Journal of Refrigeration, 31, 1123-1133.

[6] Ecir Ug˘ur Küçüksille, [2008], Data mining techniques for thermophysical properties of refrigerants, Energy Conversion and Management.

[7] A. Kilicarslan, An experimental investigation of a different type vapor compression cascade refrigeration system, Applied Thermal Engineering, 24, 2004 2611–2626.

[8] P. K. Jain,[2007] Cascade Systems: Past, Present, and Future, ASHRAE Trans. 113(1), 245-252(DA-07-027).

[9] D. Peter,[2008] Refrigeration of the Future for Air Conditioning, Air conditioning and Refrigeration Journal, 8, , 20-27.

[10] Samant Maji, [2006]Design and Development of two stage cascade refrigeration system, Mechanical Engineering Department, IIT Delhi.

[11] Reşat Selbaş, Önder Kızılkan, Arzu Şencan,[2006,] Thermo economic optimization of subcooled and superheated vapor compression refrigeration cycle, Energy, 31, 2108–2128.

[12] Xudong Wang, Yunho Hwang, Reinhard Radermacher,[2008]Investigation of potential benefits of compressor cooling, Applied Thermal Engineering, 28, , 1791–1797.

[13] Tzong-Shinge.al, [2006], Thermodynamic analysis of optimal condensing temperature of cascade-condenser in CO2/NH3 cascade refrigeration systems, International Journal of Refrigeration, 29, , 1100-1108.

[14] S. M. Zubair, [1999] Performance evaluation of Vapour Compression System, International Journal of Refrigeration, 22, 235-243.

[15] Mukhopadhyay, [2008], CO2–C3H8 cascade refrigeration–heat pump system: Heat exchanger inventory optimization and its numerical verification, international journal of refrigeration, 31, , 1207-1213

International Journal of Theoretical and Applied Mechanics.
ISSN 0973-6085 Volume 12, Number 3 (2017) pp. 555-561
© Research India Publications
http://www.ripublication.com

Thermodynamic (Energy-exergy) Analysis of Hydrogen Liquefaction System

Devender Kumar and R.S Mishra

Mechanical Department, Delhi Technological University, Shahabad, Delhi (India)

Abstract

Hydrogen is rich source of energy but its properties in gaseous state cannot be used efficiently but at the liquid state in can be used in various application but high cost of liquefaction of hydrogen cryogenic system and less efficiency component turned the cryogenic science toward optimization. In this paper First law analysis in terms of coefficient of performance and second law analysis in terms of exergetic efficiency in the system and exergy destruction ratio for finding irreversibility in the system and its components on yield base parameters have been presented for reducing the cost of whole system. A Numerical computational analysis is carried out for hydrogen liquefaction system using ambient condition and compressor pressure at 15 bar as input parameters and second law efficiency (exergetic efficiency) of hydrogen liquefaction system is obtained as 19.82 % along with first law efficiency in terms of coefficient of performance (COP) is 0.9746, It is also observed that both thermodynamic performances start decreasing with further increases of compressor pressure whereas liquefaction mass ratio and total work done in the system is increases with increasing in compressor pressure.

Keywords: Thermodynamic analysis, energy-exergy analysis, Irreversibility analysis, Cryogenic

NOMENCLATURE

m = mass flow rate

m_f = Liquefaction mass

h = Enthalpy

s = Entropy

X = Dryness fraction

T = temperature

P = Pressure

$m_{LN_{evop}}$ = mass of liquid nitrogen

$\eta_{2nd\ law}$ = Second law efficiency

Ɛ = Effectiveness of heat exchanger

C = Specific heat capacity fluid or gas

W_{net} = Total Work of compression

W_c = Compressor work

INTRODUCTION

A low temperature environment is termed a cryogenic environment when the temperature range is below the point at which permanent gases begin to liquefy. Permanent gases are elements that normally exist in the gaseous state and were once believed impossible to liquefy. Lack of fossil fuels and increasing need to energy has made us to pay a special attention to replacing fossil fuels with renewable resources. Also the fossil fuels and their combustion products were causing global environment problems. These resources of energy are clean and do not pollute the environment. One these sources are hydrogen energy. Hydrogen derived from renewable energies eventually will contribute to the sustainable development of such countries [1]. Hydrogen combustion produces water vapor that does not make any pollution. Hydrogen is gas form, occupies a large volume and has low density (0.0897 kg/m3) and high pressure [2]. so the demand of dense form of hydrogen increases but liquefaction of hydrogen is a very slow process and it estimated that The fastest flow that a hydrogen liquefaction system can bear in optimum hydrogen flow is 300 lit/hr[3]. Thus we need to liquefy hydrogen for an easier transportation and safety ,it also studied that thermodynamic performance (t) is low due to existing of nitrogen pre coolers I hydrogen system [4] In hydrogen liquefier the pre-compression of feed gas has generally higher stand-alone exergy efficiency than the cooling and liquefaction sub-process. Decreased feed pressure results in generally higher power consumption but also higher exergy efficiency, and vice versa [5].Gianluca Valenti et al [6] in research show that the feed, 10 kg s^{-1}, is refrigerated in heat exchangers catalytically promoting the ortho–para conversion down to the low temperature of 20.5 K and at the high pressure of 60 bar with turbo machine expansions show 48% of second law efficiency with low power consumption of hydrogen liquefaction system.

Songwut Krasae-in et al [7] showed effect of multi-refrigerant at the hydrogen liquefaction plant. The MR system can be used to cool feed normal hydrogen gas from 25 °C to the equilibrium temperature of −193 °C with a high efficiency. The overall power consumption of the plant is reduced from 5.35 kWh/kg_{LH2}, to minimum of 2.89 kWh/kg_{LH2}. Gianluca Valenti et al [8] discuss the influence of the thermodynamic modeling of the fluid on the simulation outcomes. Various hydrogen forms (ortho hydrogen and Para hydrogen) and their mixtures (equilibrium-hydrogen and normal-hydrogen) are studied and described in his research. According to viewpoint of David O. Berstad [9] efficiency and cost is to a large extent dependent on the efficiency of the liquefier so high-efficiency hydrogen liquefier based on mixed-refrigerant (MR) pre-cooling has been developed. Based on his models, a reduction in power consumption obtain in the range of 45–48 % . Akihiro Nakano et al [10] proposed a simple estimation method for the liquefaction rate and confirmed that the estimation method well explained the experimental result. A small-scale hydrogen liquefier with a two-stage 10 K Gifford–McMahon cycle (GM) refrigerator is confirmed the estimation method for predicting the liquefaction rate.

2.0 THERMODYNAMIC ANALYSIS OF HYDROGEN LIQUEFACTION SYSTEM:

The design is quite critical at low temperatures due to changes in thermo physical properties of hydrogen gas. Fig 1 showed the block diagram hydrogen liquefaction system in which liquefy nitrogen chamber is introduced to reduce the further temperature up to the critical temperature before J-T which is required to liquefy the hydrogen gas.

“Compressor”

$$Wc_{ideal} = m_2 * R * T_1 * ln * (\frac{P_2}{P_1}) \quad (1)$$

$$-Wc = (h_1 - h_2) - T_0 * (s_1 - s_2) \quad (2)$$

"Heat Exchanger A”

$$m_a * h_2 + m_4 * h_{13} = h_3 * m_a + m_4 * h_{14} \quad (3)$$

$$\varepsilon = \frac{m_a*(T_2-T_3)}{m_a*T_2-T_{13}*m_{13}} \quad (4)$$

$$P_3 = P_2 \quad (5)$$

"Heat Exchanger B"

$$m_b * h_2 + m_{LN_{evop}} * h_7 = h_4 * m_b + h_8 * m_{LN_{evop}} \quad (6)$$

$$\varepsilon = \frac{m_b*(T_2-T_4)}{m_b*T_2-T_7*m_{LN_{evop}}} \quad (7)$$

"Mixing of two different tem helium"

$$m_3 * c_p * (T_3 - T_5) = m_b * c_p * (T_5 - T_4) \quad (8)$$

"Heat Exchanger C"

$$Q_{LN_{evop}} = m_2 * (h_5 - h_6) \quad (9)$$

$$m_{LN_{evop}} * L_{LN} = Q_{LN_{evop}} \quad (10)$$

$$m_2 * h_5 + m_{LN} * h_{LN} = m_{LN_{evop}} * h_7 + m_2 * h_6 \quad (11)$$

"Heat Exchanger D"

$$m_2 * h_6 + m_4 * h_{12} = h_9 * m_2 + m_4 * h_{13} \quad (12)$$

$$\varepsilon = \frac{m_2 * (T_6 - T_9)}{m_2 * T_6 - T_{12} * m_4} \quad (13)$$

"Expansion Valve"

$$h_9 = h_{10} \quad (14)$$

$$h_{10} = h_{11} + x_{10} * h_{12} \quad (15)$$

Separator"

$$m_2 * h_{10} = h_{11} * m_3 + m_4 * h_{12} \quad (16)$$

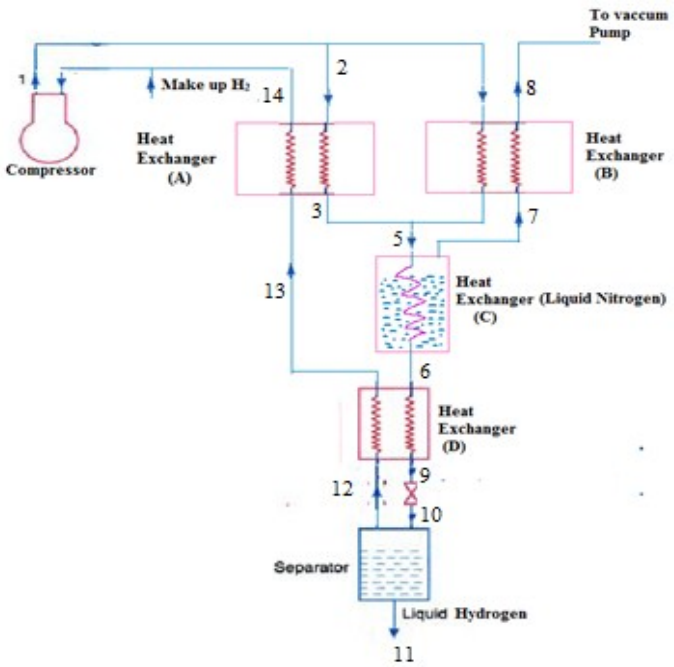


Fig 1: Block diagram of hydrogen liquefaction system

3.0 RESULT AND DISCUSSION

By computational mathematical technique various variable are noticed which are playing high role in hydrogen Liquefaction system and its optimization and to fully understand the effect of these variables on system different values are given and graphs are generated. Fig 1 show the variation of second law efficiency with respect to compressor pressure it show that as increase the compressor pressure after 12 bar the efficiency of system start decreasing whereas Fig 2 show just reverse of this and show increase in liquefaction mass of helium with increase of compressor pressure. Total work is summation of all type of work used in system like compressor work, expander work. Fig 3 shows there not much high fluctuation in variation of total work of system with increase in compressor pressure. At low pressure COP of system is quiet good as comparison with COP at high pressure. Fig 4 show COP of system start decreasing with increase in compressor pressure.

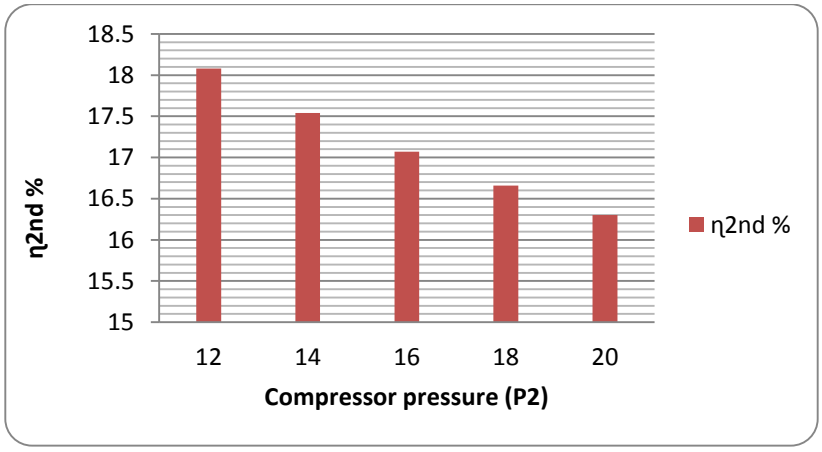


Fig:3 Variation of second law efficiency to Compressor pressure

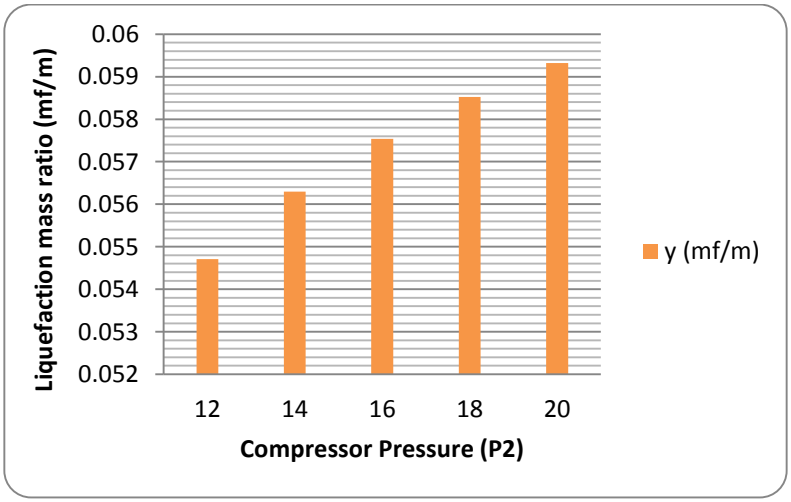


Fig:4 Variation of liquefaction mass ratio to compressor pressure

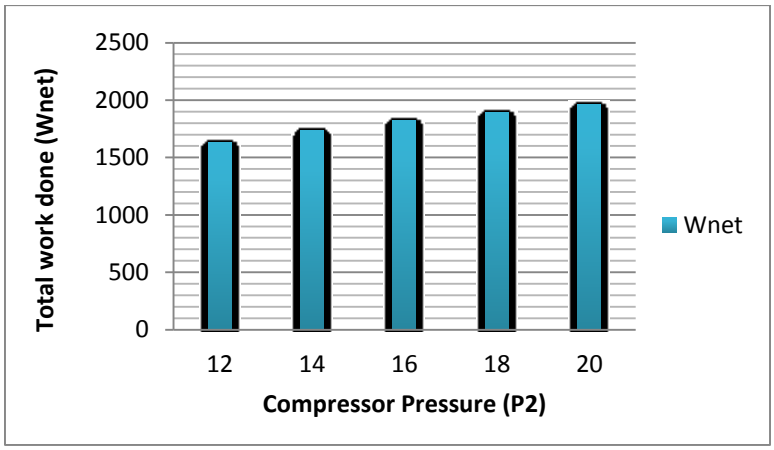


Fig:5 Variation of total work of system to Compressor pressure

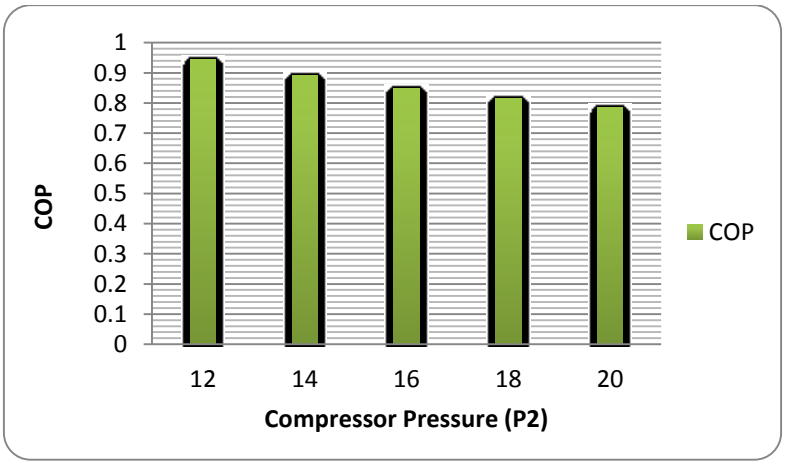


Fig:6 Variation of COP of system to Compressor pressure

CONCLUSION

1) Second law efficiency of system is 17.29 % and it start decreasing with further increases of compressor pressure
2) COP of the system is 0.8687 when input are at ambient condition and compressor pressure is 15 bar and as like second like efficiency it also start decreasing with increases of compressor pressure.
3) Liquefaction mass ratio and Total work done is increases with increase in compressor pressure.

REFERENCES

[1]. Veziroglu, T.N., Sumer Sahin, 21st Century's energy: Hydrogen energy system, Energy Conversion and Management, 49 (7), p. 1820-1831 (2008).

[2]. Aceves, S.M. , Martinez-Frias, J., Garcia-Villazana, O., Analytical and experimental evaluation of insulated pressure vessels for cryogenic hydrogen storage, International Journal of Hydrogen Energy, 25 (11), p. 1075-1085 (2000).

[3] J. Sargolzaei and Saghatoleslami "Design of cryogenic system for liquefaction of hydrogen"ICHEC 2009

[4]. Ludwig, E.E., Applied Process design for chemical and petrochemical plants, Gulf Publishing Company, 4 edition (2007).

[5] David O. Berstad, Jacob H. Stang and Petter Neksa "Comparison criteria for large-scale hydrogen liquefaction processes" International Journal of Hydrogen Energy .Volume 34, Issue 3, February 2009, Pages 1560–1568

[6] Gianluca Valenti and Ennio Macchi, "Proposal of an innovative, high-efficiency, large-scale hydrogen liquefier" International Journal of Hydrogen Energy, Volume 33, Issue 12, June 2008, Pages 3116–3121.

[7] Songwut Krasae-in ,Jacob H. Stang and Petter Neksa, "Simulation on a proposed large-scale liquid hydrogen plant using a multi-component refrigerant refrigeration system" International Journal of Hydrogen Energy, Volume 35, Issue 22, November 2010, Pages 12531–12544.

[8] Gianluca Valenti,Ennio Macchi and Samuele Brioschi, "The influence of the thermodynamic model of equilibrium-hydrogen on the simulation of its liquefaction" International Journal of Hydrogen Energy, Volume 37, Issue 14, July 2012, Pages 10779–10788.

[9] David O. Berstad, Jacob H. Stang and Petter Neksa, "Large-scale hydrogen liquefier utilising mixed-refrigerant pre-cooling", International Journal of Hydrogen Energy, Volume 35, Issue 10, May 2010, Pages 4512–4523.

[10] Akihiro Nakano, Tetsuhiko Maeda, Hiroshi Ito,Masao Masuda, Yoshiaki Kawakami, Atsushi Kato, Manabu Tange,Toru Takahashi and Masahiro Matsuo, "Small-scale hydrogen liquefaction with a two-stage Gifford–McMahon cycle refrigerator" , International Journal of Hydrogen Energy, Volume 35, Issue 17, September 2010, Pages 9088–9094

International Journal of Theoretical and Applied Mechanics.
ISSN 0973-6085 Volume 12, Number 3 (2017) pp. 563-570
© Research India Publications
http://www.ripublication.com

Thermodynamic (Exergy-energy) Analysis of a Low Pressure Improved Claude Cycle Using Control Volume Technique For Liquefaction of Gases

Devender Kumar [1]* and R.S Mishra [2]

Mechanical Department, Delhi Technological University, Shahabad, Delhi

Abstract

Comprehensive energy and exergy study of improved Claude system to liquefy various gases is carried out in this paper. Numerical computational technique is used to find out the effect of various properties variables like temperature, pressure on system performance and exergy destruction in system and their sub-components It was observed that the inlet variables like pressure temperature and intermediate mass ratio respectively are ranges from 3 to 6 bar pressure,280 to 290 K temperature and 0.5-07 mass ratio respectively desirable for high liquefaction rate and second law efficiency in improved low pressure Claude cycle.

Keywords: Thermodynamics analysis, Energy-Exergy analysis, first and second law analysis

NOMENCLATURE

m= Total mass of gas

m_f=liqified mass of gas

h=Enthalpy

s=Entropy

X=Dryness fraction

T=temperature

P=Pressure

η_{comp}=Efficiency of compressor (80%)

$\eta_{expander}$= Efficiency of expander (80%)

$\eta_{2nd\ law}$=Second law efficiency

Ɛ=Effectiveness of heat exchanger (80%)

C=Specific heat capacity fluid or gas

W_t=Work of reversible isothermal compression

W_c=Shaft work to compressor per unit mass

R=Universal gas constant

W_{net} =Net work done in system

INTRODUCTION

It's a natural phenomenon that heat flow from high temperature to low temperature and the reverse process without any aid or external work is impossible and if so it just the violation of second law of thermodynamics A device which is which act as intermediate device is called refrigerator. The difference between a refrigeration and cryogenics systems lies in the achievable temperature with the dividing line being of -100ºF or -74ºC [1]. Now a day the process industries are faced with an increasingly competitive environment, ever changing market conditions and government regulations. Yet they a still have to increase productivity and profitability. In order to have a means of comparison of liquefaction systems through the figure of merit and exergy efficiency. Most of system is ideal in the thermodynamic sense, but it is not ideal as far as practical system is concerned. The perfect cycle in thermodynamics is the Carnot cycle [2]. Cryogenic process to liquefy air which is further extent to extract various particular gas like oxygen, nitrogen, feron etc. Always various analyses is done to identify the loop hole of process and to rectify it to their upper level. electro caloric cooling is a transiting to new cooling principle's is critical and one of the most promising alternatives may be [3].Various particular part are taken under study to increase overall performance of cryogenic system e.g. A good exergetic design of a heat exchanger would allow for an increase in the global efficiency of the process, by defining a thermodynamic cycle in which the exergetic losses would be limited [4] Apart from this other parts like expander ,mass ratio and different input variables are considered to improve cryogenic systems

2.0 THERMAL ANALYSIS OF IMPROVED CLAUDE SYSTEM FOR LIQUEFACTION OF GASES:

Compressor work

$$\eta_c = \frac{W_t}{W_{comp}} \tag{1}$$

$$W_t = mRT\ln\frac{P_2}{P_1} \tag{2}$$

$$-W_c = m * (T_1 * (s_1 - s_2) - (h_1 - h_2)) \quad (3)$$

$$W_{reversible} = W_{actual} - T_0 s_{gen} \quad (4)$$

$$W_{net} = W_c + W_e \quad (5)$$

Expander

$$\frac{T_8}{T_3} = \left(\frac{P_8}{P_3}\right)^{\left(\frac{Y-1}{Y}\right)\eta_{expander}} \quad (6)$$

$$W_e = m_e * h_3 - m_e * h_e \quad (7)$$

"Control volume except compressor"

$$m * h_2 = W_e + (m - m_f) * h_1 + h_f * m_f \quad (8)$$

$$y = m_f/m \quad (9)$$

"Work done per mass of gas"

$$z = -W_{net}/m \quad (10)$$

"Work done per mass of liquefaction gas"

$$t = -W_{net}/m_f \quad (11)$$

Coefficient of performance of system

$$COP = ((h_1 - h_f)/W_{net}) \quad (12)$$

Second law analysis

$$\eta_{2nd\ law} = ((((h_f - h_1) - T_0 * (s_f - s_1))/(W_{net}) * m_f) * 100) \quad (13)$$

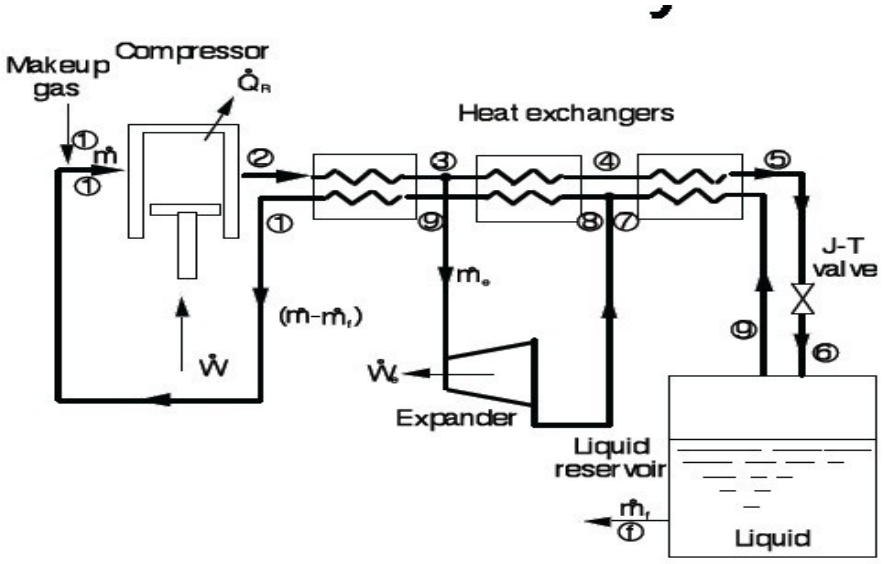


Fig 1 Block diagram of low pressure Claude system

A complete analysis of Improved Claude cycles performed with the help of numerical computation technique for various gases. Improved Claude cycles shown in Fig 1 is taken for analysis. Improved Claude system is almost same as simple Claude system except arrangement of first expander .In Improved Claude system expander is situated in between the first and second heat exchanger other than this it also consist a compressor, expander, two heat exchangers with throttle valve and separator. The fluid which has to liquefy first fed to compressor in its gaseous form at atmospheric pressure and temperature which circulate from all system and in last fractional mass of total mass get liquefied and remaining again fed in system with additional mass to recirculate in system again. Various results are drawn for particular inlet temperature, pressure and intermediate pressure for low pressure side of expander for different gases such oxygen ,argon, methane ,air ,fluorine and nitrogen are considered for study.

RESULTS AND DISCUSSION

Various results are drawn on the basis of numerical equations of system. In fig 2 variation of liquefaction temperature with inlet pressure as we increases the pressure Liquefaction temperature rises but after crossing 10 bar the increment in liquefaction temperature is start reducing and its slope with inlet pressure start become straighten. Fig 3 show fall of liquefaction mass with increase of inlet temperature and it also show that at 330 k the liquefied mass of methane and argon is same. Fig 4 shows decreases in liquefaction mass with increase of inlet pressure. Fig 5-7 show variation in second law efficiency with Intel temperature, intermediate mass and inlet pressure respectively. Graph analysis of these 5-7 fig shows that second law efficiency is decreases with increase of inlet temperature while with increases of intermediate mass

second law efficiency increases whereas it again decreases with increase of inlet pressure. Fig 8 -9 show variation in COP of system with inlet pressure and temperature. They show that increase in pressure is beneficial for system and COP of system is increases with increase in inlet pressure while its increment with increase in inlet temperature is very less. From above graph study it determined that increment and decrement in a very concern range of various variables is good for optimization of Improved Claude system

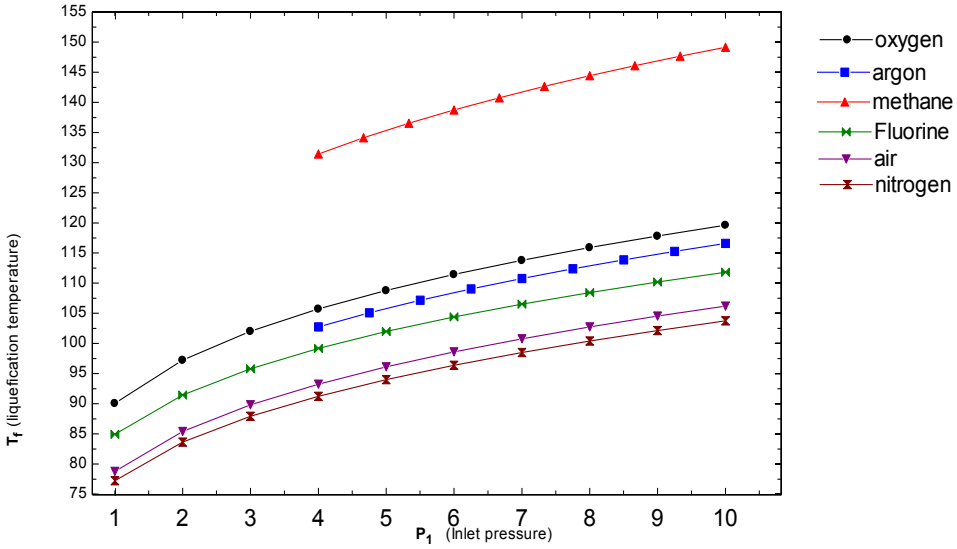


Fig: 2 Variation of liquefaction temperature with inlet Pressure

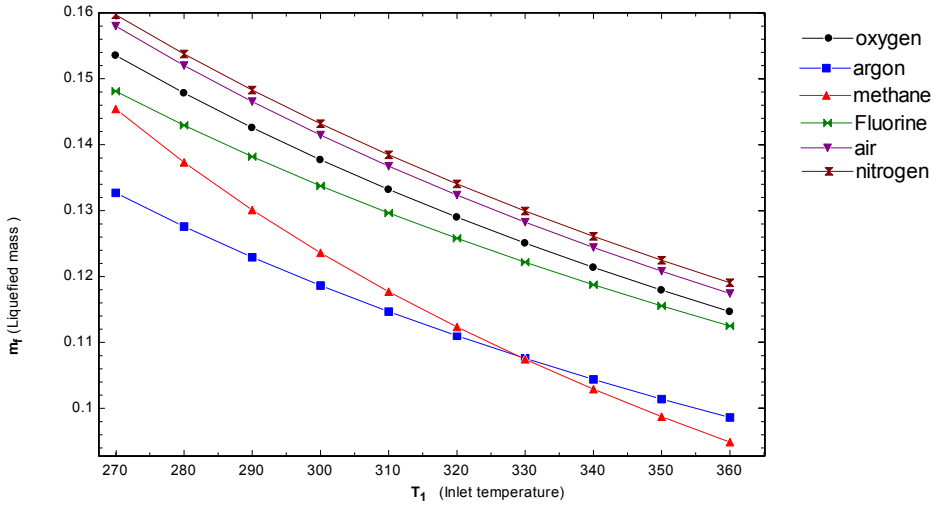


Fig:3 Variation of liquefied mass with inlet temperature

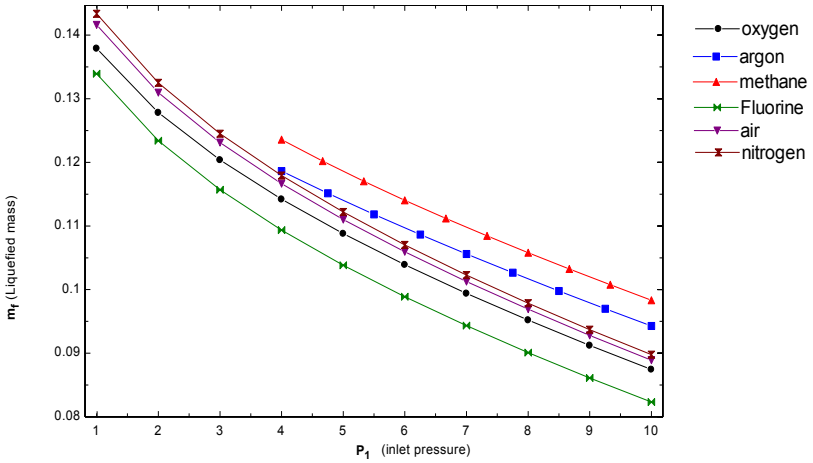


Fig:4 Variation of liquefied mass with inlet Pressure

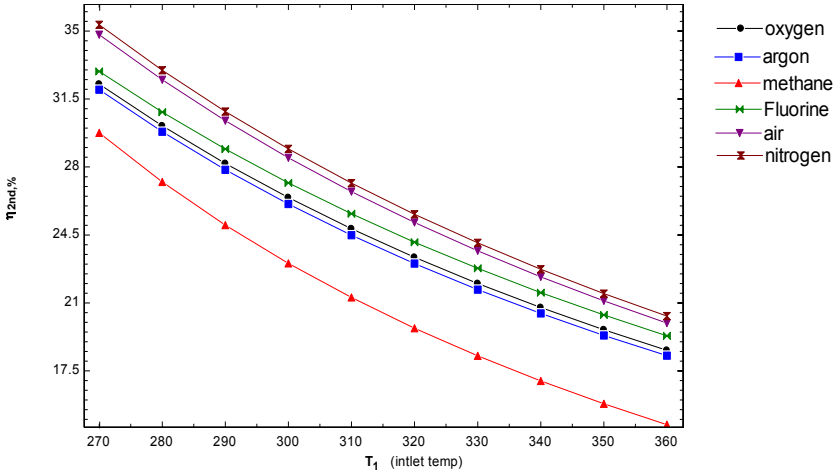


Fig:5 Variation of 2nd law efficiency with inlet temperature

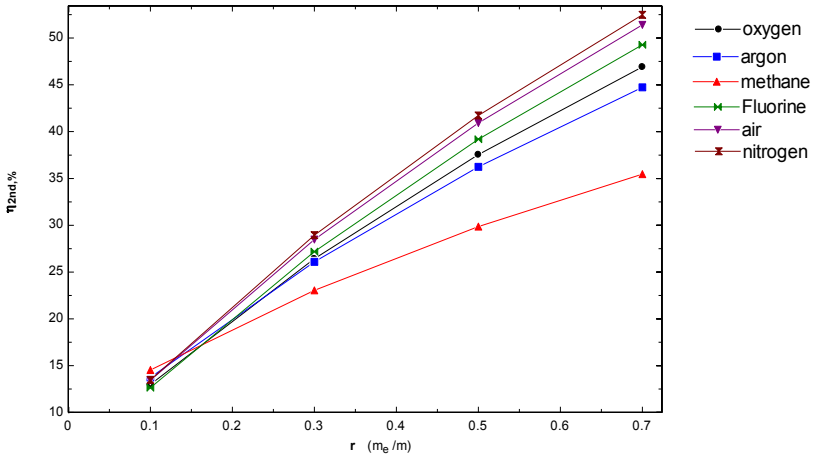


Fig:6 Variation of 2nd law efficiency with intermediate mass ratio

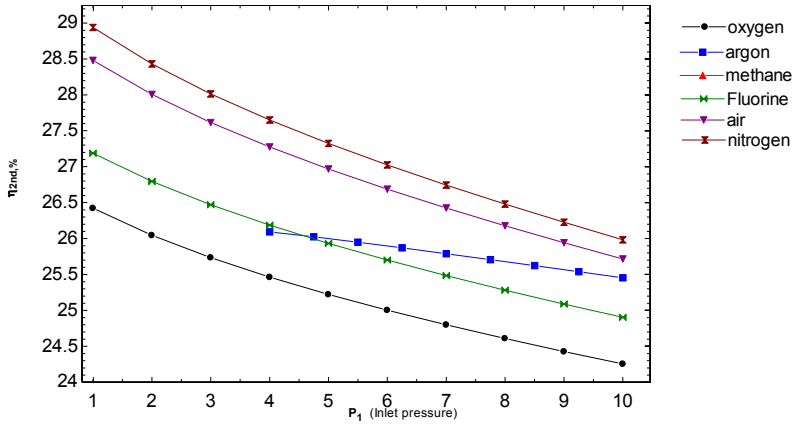


Fig:7 Variation of 2nd law efficiency with inlet Pressure

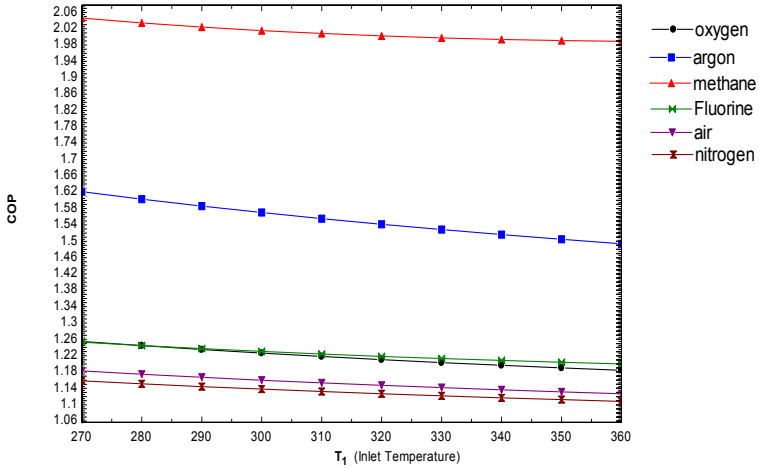


Fig:8 Variation of COP with inlet Temperature

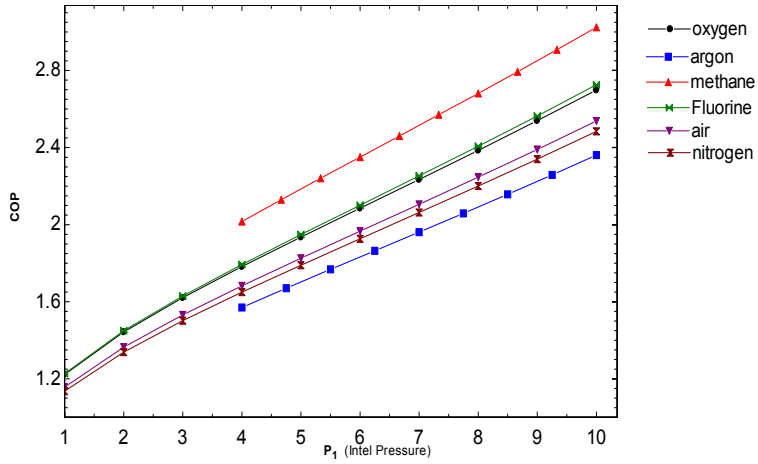


Fig:9 Variation of COP with inlet pressure

CONCLUSION

From above study following results are concluded

1) The optimum value of inlet variables like pressure temperature and intermediate mass ratio respectively ranges from 3 to 6 bar pressure,280 to 290 K temperature and 0.7 mass ratio for good result of considered variables such liquefaction mass, liquidation temperature and second law efficiency.
2) Intermediate pressure of low pressure side expander should be in minimum range for high second law efficiency.
3) Increase in inlet temperature decrease the COP, second law efficiency and output (liquefaction mass) of system.

REFERENCES

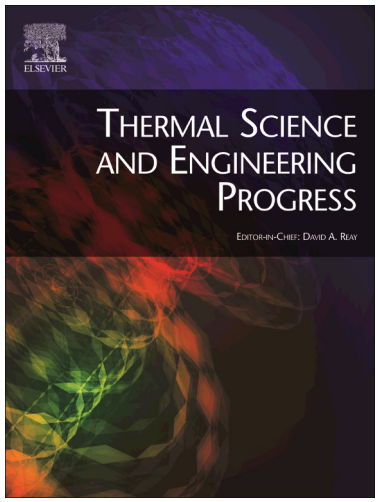
[1] A.M . Papadopoulos and C.J Koroneos "Refrigeration and Cryogenic systems" Air conditioning –Energy consumption abd environmental quality, Encyclopedia of life supports systems (ELOSS)

[2] Sunil Manohar Dash "Study of cryogenic cycles with aspen – hysys simulations" Department of Mechanical Engineering National Institute of Technology Rourkela 2008-09

[3] Yu. V. Sinyavskii, "Electrocaloric refrigerators: A promising alternative to current low-temperature apparatus" Chemical and Petroleum Engineering, June 1995, Volume 31, Issue 6, pp 295-306, ISSN 1573-8329.

[4] R. Agrawal, D.W. Woodward, Efficient cryogenic nitrogen generators: An exergy analysis, Gas Separation & Purification, Volume 5, Issue 3, September 1991, Pages 139-150, ISSN 0950-4214.

Accepted Manuscript



Thermodynamic analysis and optimization of a supercritical CO_2 regenerative recompression Brayton cycle coupled with a marine gas turbine for shipboard waste heat recovery

O.P. Sharma, S.C. Kaushik, K. Manjunath

PII: S2451-9049(17)30021-5
DOI: http://dx.doi.org/10.1016/j.tsep.2017.06.004
Reference: TSEP 36

To appear in: *Thermal Science and Engineering Progress*

Received Date: 8 February 2017
Revised Date: 31 May 2017
Accepted Date: 3 June 2017

Please cite this article as: O.P. Sharma, S.C. Kaushik, K. Manjunath, Thermodynamic analysis and optimization of a supercritical CO_2 regenerative recompression Brayton cycle coupled with a marine gas turbine for shipboard waste heat recovery, *Thermal Science and Engineering Progress* (2017), doi: http://dx.doi.org/10.1016/j.tsep.2017.06.004

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. 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.

ACCEPTED MANUSCRIPT

Thermodynamic analysis and optimization of a supercritical CO_2 regenerative recompression Brayton cycle coupled with a marine gas turbine for shipboard waste heat recovery

O P Sharma[a*], S C Kaushik[a*], K Manjunath[b]

[a] Centre for Energy Studies, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110016, India
[b] Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

Abstract

This communication presents the modelling and analysis of a waste-heat-recovery-system (WHRS) using supercritical carbon dioxide (SC-CO_2) based regenerative recompression Brayton cycle (RRCBC) for shipboard applications following energy and exergy analyses. The influence of key operating parameters such as, the gas compositions, turbine and compressor inlet temperatures, pressure drop irreversibility, pressure ratio etc., on the overall performance of the system including the exergy destruction rate has been investigated. The results show that the proposed integration improves the overall efficiency of the system by 10% while the net power is found to be increasing up to 25% of the rated power. It is also found that the topping gas turbine exhaust gas compositions and temperatures have a significant influence on the WHRS performance. The results are also found in good agreement with those already available in the published literature.

Keywords: waste heat recovery; supercritical carbon dioxide; regenerative recompression Brayton cycle; energy and exergy analysis; gas compositions.

* Corresponding author(s): kaushik@ces.iitd.ac.in; opsharma.iitd@gmail.com

ACCEPTED MANUSCRIPT

1. Introduction

Due to the increasing consumption of fossil-fuel, soaring cost of crude and associated environmental issues such as environmental pollution, depletion of ozone layer and global warming, the commercial and military naval fleet are required to explore the sustainable energy technology for the integration with the existing conventional systems. The modification in the existing shipboard systems can be undertaken in different ways such as the waste heat recovery from the exhaust of the gas turbine and its utilization for heating and power generation applications. This will also improve the overall performance of shipboard systems which has become a focused area for naval fleets in the recent time from the environmental as well as economical point of views [1]. Due to some unique features, the supercritical carbon dioxide (SC-CO_2) power cycles are being extensively studied as the potential candidate for improved performance of the thermal energy conversion systems such as, advanced nuclear power, solar thermal power, automotive and geothermal power. These cycles principally use CO_2 as the working fluid and operate entirely in its supercritical phase i.e. region above the critical point which was originally proposed and studied by Feher [2-3], who obtained a few significant observations. For example, supercritical CO_2 cycle offers various characteristics desirable in a practical application such as high thermal efficiency, low volume-to-power ratio, no blade erosion in the turbine, no cavitation in the pump, single stage turbomachinery, single phase fluid in the heat rejection process, and insensitivity to compression efficiency.

Due to some unique features, such as the compactness, high power density and modularity, the supercritical CO_2 engines can be explored for number of practical applications in space, land, and sea based power generation systems. Angelino [4] studied the thermodynamic properties of supercritical CO_2 and concluded that due to its behaviour close to a real gas in the vicinity of the Andrews curve, it can be used as the potential working fluid in the air standard cycles. He also found that the supercritical CO_2 can achieve substantially better performance than that of the conventional working fluid at higher (moderate) inlet pressure and temperature of the turbine. Combs [5] investigated the performance of a supercritical CO_2 engine for propulsion power in a naval ship application using basic thermodynamics approach. He reported that the supercritical CO_2 cycles can achieve higher power and efficiency with significant fuel savings. Marron [6] evaluated the feasibility of simple and combined cycle engines with waste heat recovery system based on

ACCEPTED MANUSCRIPT

life cycle cost and its impact on the shipboard machinery arrangement. He also proposed a steam based system named as Rankine cycle energy recovery (RACER) system, wherein a steam cycle, driven by the exhaust waste heat is employed as bottoming cycle to produce power. However, RACER was found to be a complex system due to large size and weight of components besides, the higher maintenance of important parts such as, the heat recovery unit, condenser, connecting tubes, turbine, etc. which has restricted its application in naval shipboard applications.

Dostal et al. [7-8] investigated different layouts of the supercritical CO_2 Brayton cycles for a gas-cooled, modular fast nuclear reactor of 600 MW and compared their thermo-economic performance with that of the conventional steam and helium (Brayton) cycles. Using the basic thermodynamic approach, these studies reported a few significant observations. For example, a Brayton cycle using supercritical CO_2, due to its real gas behaviour, can achieve higher thermal efficiencies when operated at medium temperature (500 – 700 ^{0}C) as compared to the superheated steam and helium (Brayton) cycles. Dostal [9] also evaluated the supercritical CO_2 Brayton cycles for advanced nuclear reactor for power generation applications. He further indicated that these systems are ten times more compact and are much simpler in design than those of the conventional cycles. He also observed that for a set of operating condition (i.e. turbine inlet temperature of 550 ^{0}C with 200 bar pressure), the thermal efficiency of supercritical CO_2 recompression Brayton cycle was found to be in the range of 46-50%, while in case of helium Brayton cycle the same range of thermal efficiency can be achieved but the turbine inlet temperature of 800 ^{0}C is required, which obviously has the higher operating and maintenance cost.

Moroz et al. [10] studied the performance of a supercritical CO_2 and a conventional steam cycle for exhaust heat recovery for a shipboard gas turbine of 53.26 MW and reported a few significant observations such as using a supercritical CO_2 recompression cycle, the net power of the gas turbine was found to be increased by about 30% as compared about 26% in case of steam cycle. Bella [11, 12] analysed the supercritical CO_2 Brayton cycle with and without reheat options, for exhaust heat recovery for a shipboard gas turbine incorporating the thermo-electric generator (TEG) and reported an improvement in power and fuel savings by 22% and 13%, respectively. However, due to constraints of high cost and low efficiency, the practical feasibility of TEG system was found to be challenging at the moment and will require more R&D work in this direction.

ACCEPTED MANUSCRIPT

Ishiyama et al. [13] studied the steam, helium and supercritical CO_2 cycles for power generation in a prototype nuclear fusion reactor of total 2910 MW having a rector outlet temperature of around 500 ^{0}C. Using the basic thermodynamic approach, they obtained a few important observations using steam, helium and supercritical CO_2 as the working fluid. For example, their cycle efficiencies, at a heat source temperature of 480 ^{0}C, were obtained of the order of 40%, 34% and 42%, respectively, and also observed the high compactness of the CO_2 system over the conventional systems. Sarkar [14] studied the supercritical CO_2 regenerative recompression cycle for an advanced nuclear reactor as a heat source with a reactor outlet temperature in the range of 550-800 ^{0}C using the first as well as the second law of thermodynamics. He also presented a few important observations about the energetic and exergetic efficiencies at a turbine inlet temperature of 550 ^{0}C. The maximum rate of exergy destruction was observed in the heat source (reactor) followed by the high temperature recuperator (HTR), while it was found to be reverse in the case of main compressor.

Moisseytsev and Sienicki [15] studied different supercritical CO_2 Brayton cycles for the advanced burner test reactor (ABTR) and sodium-cooled fast reactor (SFR) having a reactor outlet temperature of 510 ^{0}C using the basic thermodynamic approach. They reported that the thermal efficiency for the recompression cycles is found to be the highest among all the cycles given above. For example, at turbine inlet temperature of 470 ^{0}C, their thermal efficiencies were obtained of the order of 40% which is found to be significant as compared to those already reported in the literature. Dyreby et al. [16, 17] studied the supercritical CO_2 Brayton cycles using semi-empirical formulations for each components and obtained a few important observations while conductance model was used for the recuperator. The thermal efficiency was found to be 47% and 49%, for the simple recuperated and the recompression CO_2 cycles, respectively. Pham et al. [18] studied the different configurations of supercritical CO_2 recompression cycle including reheated, intercooled and partial cooling variants for a small modular reactor and a SFR, using energy and exergy analysis. They also obtained a maximum cycle efficiency of 44% while the operating turbine inlet temperature of 515 ^{0}C was used, whereas the compressor inlet temperature was kept as 35 ^{0}C. Zhao et al. [19] studied the supercritical CO_2 recuperated and recompression Brayton cycles for advanced nuclear power application having a reactor outlet temperature of 750 ^{0}C, using a basic thermo-economic approach. They also achieved

ACCEPTED MANUSCRIPT

the thermal efficiency of the recuperated and the recompression variants of supercritical CO_2 Brayton cycle as of 54% and 63%, respectively.

Few researchers have focussed on possible use of supercritical CO_2 cycles for power generation in concentrated solar power (CSP) based systems for different applications [20-24]. For example, Zhang et al. [20] studied the supercritical CO_2 Rankine cycle driven by solar energy for combined power and heat generation, using the basic thermodynamic approach and obtained a few important observations. Garg et al. [21] studied three variants (i.e. subcritical, transcritical and supercritical) of CO_2 Brayton cycles, with entirely gaseous phase operation of CO_2, using CSP involving a typical heat source temperature of 600 ^{0}C. They obtained a few important results, using the supercritical cycle such as, the thermal efficiency of 30% even at a lower turbine inlet temperature of 547 ^{0}C, whereas a conventional subcritical CO_2 cycle requires a higher turbine inlet temperature of 705 ^{0}C to achieve the same thermal efficiency. They further reported that the supercritical cycle produced approximately 5 times higher specific power output than that of a conventional subcritical CO_2 cycle, for the same optimum pressure ratios at any given turbine inlet temperature.

Turchi et al. [22] explored three variants of supercritical CO_2 recompression (Brayton) cycles for CSP applications considering heat source temperature in the range of 500-850 ^{0}C and obtained an important observation that the thermal efficiency of more than 50% can be achieved by the recompression cycles and such systems can also accommodate dry cooling option. Similarly, Neises and Turchi [23] carried out the comparative study of supercritical CO_2 power cycle for different configurations. They found that the recompression cycle can achieve the highest efficiency among the three different configurations, studied by them. Padilla et al. [24] studied the supercritical CO_2 Brayton cycles integrated solar central receivers using the exergy analysis. They reported that the recompression cycle with main compression intercooling was found to have achieved the best thermal performance with exergy efficiency of 55.2 % for the turbine inlet temperature of 850 ^{0}C.

Chen et al. [25] studied four variants of CO_2 Brayton cycles in an automotive application to reduce specific fuel consumption using the basic thermodynamic approach. They obtained important results for exergetic performance and found that the transcritical CO_2 power cycle has a better thermal efficiency as compared to the supercritical CO_2 power

ACCEPTED MANUSCRIPT

cycle for the case of low grade small-scale energy recovery. Sanchez et al. [26] studied air and supercritical CO_2 Brayton cycles as a bottoming cycle for enhancing the efficiency of molten carbonate fuel cells based power generation system. The study reported that using supercritical CO_2 bottoming cycle, the overall thermal efficiency of the combined (hybrid) system obtained was found to be of 60% which represents a 10% increase as compared to the reference fuel cell system. Clementoni et al. [27] investigated the performance of an integral system test (IST) facility designed to produce 100 kWe power using supercritical CO_2 recuperated Brayton cycle. Lambruschini et al. [28] carried out a dynamic simulation of a 10 MWe closed-loop supercritical CO_2 recompression Brayton cycle model. Using a MATLAB-Simulink software, they reported that the thermal efficiency of the power plant, at a turbine inlet temperature of 700 ^{0}C and cycle pressure ratio of 2.8 was obtained as of 47%, while a 10% decrease in the heat input to the cycle resulted in the lowering of thermal efficiency by 6%. Ahn et al. [29] reviewed the technology and R&D status of supercritical CO_2 power cycles and reported number of important observations along with the brief outcomes of the available literature in a very systematic way.

The waste heat recovery system comprises of a heat recovery steam generator (HRSG), in case of steam Rankine system while it comprises of a heat recovery heat exchanger (HRHE) in case of supercritical CO_2 Brayton cycle. These systems generate steam or hot vapour of working fluid by utilizing the energy available in the exhaust of the topping gas turbine. Also, from an economic standpoint their performance is very important to the efficient and effective operation of advanced power conversion systems as each additional unit of steam or vapour generated represents additional power generation or process heat for application. Thus, the design and optimization of an HRSG/HRHE plays a vital role in optimizing the system performance to maximize efficiency, heat recovery and net power output for a given WHRS [30]. Following Ganapathy [30], other researchers have focussed on enhancing the heat transfer performance through different techniques using different approaches such as, the effect of optimal Reynolds number for the forced convection in a helical coiled tube [31], optimization of the pin-fins and the plate and fin heat exchangers [32-33]; optimum sizing and adiabatic spacing of heat sources cooled by laminar forced convection of a fluid [34]; and use of different types of shapes and geometries of the heat exchangers [35-38]. These studies have reported some fruitful results in the optimization of thermal processes for maximising heat transfer through such modified techniques.

ACCEPTED MANUSCRIPT

Kaushik and Tyagi [39] carried out the parametric study of an irreversible regenerative Brayton cycle with finite heat capacity of external reservoirs and obtained some important results for a better performance using the concept of finite time thermodynamics. Kaushik et al. [40] carried out a detailed study on the irreversible regenerative Brayton cycle with isothermal heat addition with the finite heat capacity of the external reservoirs. They further reported that there was a significant improvement of above 15%, in the performance of the Brayton cycle, with isothermal heat addition over the conventional Brayton cycle. They also concluded that the inlet temperature for the isobaric heat addition should be less than the inlet temperature of the isothermal heat addition for better performance because the heat addition in the former increases the quantity of energy, while in the latter, it increases the quality of energy (exergy) as the isothermal heat addition increases the kinetic energy rather than the thermal energy unlike the former source. Tyagi et al. [41] carried out a detailed parametric study of a regenerative-intercooled-reheat Brayton cycle and reported that the maximum power output was found to be an increasing function of the effectiveness of different heat exchangers, heat capacitance rates in the external reservoirs, component efficiencies while it was found to be a decreasing function of the working fluid heat capacitance rate. Further it was reported that there have been the optimal values of the intercooling, reheat and cycle pressure ratios and the turbine outlet temperature at which the cycle attained the maximum power output and maximum thermodynamic efficiency.

The earlier studies mentioned so far have been based on the Brayton power cycles using supercritical CO_2 as the working fluid for nuclear power, concentrated solar power, automotive, fuel cell application and the finite time thermodynamics. The most of studies mentioned above, have been based solely on the first law analysis, also called energy analysis, which provides information into how efficiently energy is converted, while the exergetic efficiency based on the second law analysis represents the energy “quality”. Unlike energy, exergy is destroyed by irreversibilities (entropy generation) such as friction, heat loss or chemical irreversibilities and cannot be recovered like energy. More specifically, the exergy analysis is concerned with the degradation of energy during a process, the entropy generation, and the lost-opportunities-to-do-work [42]. The second law analysis provides an effective technique for measuring and optimizing performance of a thermal system by accounting for the energy quality and for thermal systems, it is widely

ACCEPTED MANUSCRIPT

gaining acceptance over traditional energy methods in both industry and academia as it is developed into a set of standards for measuring the performance [43]. Butcher and Reddy [44] studied the second law analysis of a waste heat recovery based power generation system for steam based waste heat recovery power generation system. Also the different versions of power cycles with and without modifications are given by Moran et al. [45], whereas the basics of processes and the operation of cycles along with the empirical relations based on energy and exergy approach were given by Cengel and Boles [46] with the usage of some software like EES [47]. On the other hand, Span and Wagner [48], proposed a new equation of state for carbon dioxide covering the fluid region from the triple point temperature to higher temperature and pressures as high as 1100K and 8000 bar, respectively.

However, it is observed from the literature survey that majority of work for supercritical regenerative and recompression CO_2 power cycles is limited to the first law of thermodynamics but the second law analysis based on the exergy (availability) is scant and need to be looked into for the better understanding of irreversibilities associated with these system for shipboard applications.

Thus in the present work, the authors have made efforts for the modelling and analysis of waste-heat-recovery-system using supercritical carbon dioxide regenerative recompression Brayton cycle for shipboard applications based on energy and exergy analyses. Besides, the influence of key operating parameters such as, gas composition, gas inlet temperature, turbine and compressor inlet temperatures, pressure drop irreversibility, pressure ratio etc., on the overall performance of the system including the exergy destruction rate has been investigated in detail using a typical database, available somewhere else [49].

2. System description and analysis

The supercritical CO_2 regenerative recompression Brayton cycle (RRCBC) is basically a closed-loop Brayton (power) cycle using CO_2 as the working fluid. Since the operating pressures and temperatures of this Brayton cycle are maintained above the critical point, it is termed as supercritical CO_2 Brayton cycle. The RRCBC consists of seven main components, turbines, compressors, heat exchangers, heat recovery heat exchanger, precooler, etc. as shown in Figs.1, 2(a) and 2(b). The exhaust of the topping cycle at state

ACCEPTED MANUSCRIPT

point g_{in} enters the heat-recovery-heat-exchanger (HRHE) and transfers the heat during process g_{in}-g_{out} to the supercritical CO_2 as the working fluid (bottoming cycle) which gets heated to state point 1 during the heat addition process 8-1. The superheated CO_2 at state point 1 enters the turbine and expands to state point 2, while delivering power during expansion process 1-2. Further, the exhaust of the turbine enters the high temperature regenerator (HTR) and cools down to state point 3, while transferring the heat to the exhaust of the second compressor during process 7-8, as shown in Fig.1. Subsequently, the excess heat of the CO_2 is utilized in the low temperature regenerator (LTR) during process 3-4, while this heat is taken away by the exhaust of the first compressor (called as main compressor or MC) during process 6-7. The stream at state point 4 partly enters the second compressor (called as recompression compressor or RC) and partly precooled in a precooler to state point 5 before entering the main compressor. At state point 7, both the streams get mixed together and enters the high temperature regenerator and gets internally heated to state point 8 before entering the heat recovery heat exchanger where it finally, gets superheated to state point 1 thereby, completing the cycle.

As can be seen from Fig.1, the waste heat recovery system (WHRS) is basically a combined cycle power plant comprising of a shipboard gas turbine (topping engine) and the SC-CO_2 based regenerative recompression Brayton cycle (bottoming cycle) to recover energy from the exhaust for augmentation of power. The broad specifications of the topping gas turbine are indicated in Table 1. As the exhaust from the topping gas turbine is passed through the HRHE, the useful energy available in the exhaust is transferred to the working fluid (SC-CO_2) and finally, utilised for the power production in the bottoming cycle. The energetic and exergetic analysis of RRCBC and WHRS are presented in the succeeding sub-paragraphs, while the salient assumptions are given as below:

2.1 Assumptions

In order to facilitate the mathematical modelling and simulation of RRCBC and WHRS for the energy and exergy analysis, certain assumptions such as the steady-state steady-flow (SSSF) and closed loop operation of the cycle; the adiabatic boundaries with a single control volume for the cycle components, negligible change in the kinetic and the potential energy, negligible heat transfer with the ambient in all components except precooler, no leakage of working fluid, no pressure loss and the heat loss in pipelines to the ambient, etc.

ACCEPTED MANUSCRIPT

are considered. On the other hand, the exergy carried out by the outgoing streams, the gas outlet from the HRHE to the funnel (atmosphere) and the cooling water outlet stream from the precooler to the ambient (sea), were not considered since they have not been utilized further.

In order to assess the impact of variation in heat exchanger's pressure drop on the cycle performance, a non-dimensional relative pressure drop fraction, $\delta\bar{P}_{ij}$ across a stream ($i-j$) of the working fluid, ($\delta\bar{P}_{ij}=\Delta P/P_i$), where ΔP = (P_i-P_j), has been defined, and can be expressed for various heat exchangers, as below:

$$\delta\bar{P}_{81(RRCBC)}=\left(P_8-P_1\right)/P_8 \quad \text{: (across the } CO_2 \text{ stream of the HRHE)} \tag{1}$$

$$\delta\bar{P}_{23(RRCBC)}=\left(P_2-P_3\right)/P_2 \quad \text{: (across the hot stream of the HTR)} \tag{2}$$

$$\delta\bar{P}_{34(RRCBC)}=\left(P_3-P_4\right)/P_3 \quad \text{: (across the hot stream of the LTR)} \tag{3}$$

$$\delta\bar{P}_{45(RRCBC)}=\left(P_4-P_5\right)/P_4 \quad \text{: (across } CO_2 \text{ stream of the precooler)} \tag{4}$$

$$\delta\bar{P}_{67(RRCBC)}=\left(P_6-P_7\right)/P_6 \quad \text{: (across the cold stream of the LTR)} \tag{5}$$

$$\delta\bar{P}_{78(RRCBC)}=\left(P_7-P_8\right)/P_7 \quad \text{: (across the cold stream of the HTR)} \tag{6}$$

The turbine and the compressors efficiencies are defined as below [41-42]:

$$\eta_{isen,turbine}=\left(h_1-h_2\right)/\left(h_1-h_{2s}\right) \tag{7}$$

$$\eta_{isen,MC}=\left(h_{6s}-h_5\right)/\left(h_6-h_5\right) \tag{8}$$

$$\eta_{isen,RC}=\left(h_{7s}-h_4\right)/\left(h_7-h_4\right) \tag{9}$$

where h_{2s} is the enthalpy at turbine outlet state (2s), h_{6s} = enthalpy at main-compressor outlet state (6s) and h_{7s} = enthalpy at re-compression compressor outlet state (7s). The heat exchangers effectiveness can be expressed as [9, 42]:

$$\varepsilon_{LTR}=\left(h_3-h_4\right)/\left(h_3-h_{6'}\right) \tag{10}$$

$$\varepsilon_{HTR}=\left(h_2-h_3\right)/\left(h_2-h_{7'}\right) \tag{11}$$

$$\varepsilon_{HRHE}=\left(T_{g,in}-T_{g,out}\right)/\left(T_{g,in}-T_8\right) \tag{12}$$

where, $h_{6'}$ represents the enthalpy at $T=T_6$, $P=P_4$; $h_{7'}$ represents the enthalpy of the working fluid at $T=T_7$, $P=P_3$ while it is assumed that the minimum heat capacity of the

ACCEPTED MANUSCRIPT

working fluid is situated in the hot streams, (3-4) in the LTR, (2-3) in the HTR and the exhaust gas stream in the HRHE.

2.2 Energetic analysis

Based on the first law of thermodynamics, the energy balance equations for the LTR, the HTR, the precooler and the HRHE, respectively are given by [40-42]:

$$h_3 - h_4 = (1-x)(h_7 - h_6) \tag{13}$$

$$(h_2 - h_3) = (h_8 - h_7) \tag{14}$$

$$(1-x)\dot{m}_{CO_2}(h_4 - h_5) = m_w(h_{w,out} - h_{w,in}) \tag{15}$$

$$\dot{m}_{CO_2}(h_1 - h_8) = m_g(C_{p,g,in}T_{g,in} - C_{p,g,out}T_{g,out}) \tag{16}$$

Let $\dot{Q}_{in,RRCBC}$ and $\dot{Q}_{out,RRCBC}$ be the heat transfer rates to and from the cycle respectively, then [40-42]:

$$\dot{Q}_{in,RRCBC} = \dot{m}_{CO_2}(h_1 - h_8) \tag{17}$$

$$\dot{Q}_{out,RRCBC} = (1-x)\dot{m}_{CO_2}(h_4 - h_5) \tag{18}$$

Using the first law of thermodynamics, the expressions for the work across turbines and compressors and the net power output of the cycle are given by [40-42]:

$$\dot{W}_{turbine} = \dot{m}_{CO_2}(h_1 - h_2) \tag{19}$$

$$\dot{W}_{MC} = (1-x)\dot{m}_{CO_2}(h_6 - h_5) \tag{20}$$

$$\dot{W}_{RC} = x\dot{m}_{CO_2}(h_6 - h_5) \tag{21}$$

$$\dot{W}_{net} = \dot{W}_{turbine} - (\dot{W}_{MC} + W_{RC}) \tag{22}$$

From the first law of thermodynamics, the thermal efficiency or the energetic efficiency and the Back-Work ratio or BWR (ratio of the work consumed by the compressor(s) and the work generated by the turbine(s)) of the RRCBC are given as [40-42, 46]:

$$\eta_{I,RRCBC} = \dot{W}_{net} / \dot{Q}_{in,RRCBC} \tag{23}$$

$$BWR_{RRCBC} = (\dot{W}_{MC} + \dot{W}_{RC}) / W_{turbine} \tag{24}$$

2.3 Exergetic analysis

ACCEPTED MANUSCRIPT

Using the second law of thermodynamics, the general expression for specific exergy flow at each state point (i) of the cycle for SSSF (i = 1 to 8) is given as [46]:

$$e_i = \left(h_i - h_0\right) - T_0\left(s_i - s_0\right) \tag{25}$$

while the total exergy input by heat, to the cycle is given as [46]:

$$\dot{E}_{in,RRCBC} = \dot{Q}_{in,RRCBC}\left(1 - T_0 / T_{g,\,in}\right) \tag{26}$$

where, the heat source temperature which is the gas inlet temperature, $T_{g,\,in}$ (K), is assumed as constant. The rate of irreversibility for a component of the cycle is expressed as the sum of the exergy inputs to the component minus the sum of the exergy outputs from the component while the non-dimensional fraction (i_j) of irreversibility for the j^{th} component of a cycle is defined as the ratio of the irreversibility rate of the j^{th} component and the total exergy input rate of the cycle, as given below [14]:

$$i_{turbine} = \left\{\dot{m}_{CO_2}\left(e_1 - e_2\right) - \dot{W}_{turbine}\right\} / \dot{E}_{in,RRCBC} \tag{27}$$

$$i_{mc} = \left\{\dot{m}_{CO_2}\left(1 - x\right)\left(e_5 - e_6\right) + \dot{W}_{MC}\right\} / \dot{E}_{in,RRCBC} \tag{28}$$

$$i_{rc} = \left\{\dot{m}_{CO_2} x\left(e_4 - e_7\right) + \dot{W}_{RC}\right\} / \dot{E}_{in,RRCBC} \tag{29}$$

$$i_{HTR} = \left[\dot{m}_{CO_2}\left\{\left(e_2 - e_3\right) + \left(e_7 - e_8\right)\right\}\right] / \dot{E}_{in,RRCBC} \tag{30}$$

$$i_{LTR} = \left[\dot{m}_{CO_2}\left\{\left(e_3 - e_4\right) + \left(1 - x\right)\left(e_6 - e_7\right)\right\}\right] / \dot{E}_{in,RRCBC} \tag{31}$$

$$i_{precooler} = \left\{\dot{m}_{CO_2}\left(1 - x\right)\left(e_4 - e_5\right)\right\} / \dot{E}_{in,RRCBC} \tag{32}$$

$$i_{HRHE} = \left\{\dot{E}_{in,RRCBC} + \dot{m}_{CO_2}\left(e_8 - e_1\right)\right\} / E_{in,RRCBC} \tag{33}$$

Using the second law of thermodynamics, the exergetic efficiency of an energy recovery cycle is expressed as the ratio of the net exergy output to the exergy input to the bottoming cycle through the HRHE [14, 46]. Thus:

$$\eta_{II,RRCBC} = \dot{W}_{net} / \dot{E}_{in,RRCBC} = 1 - \sum_{j=1}^{8} i_{j,RRCBC} \tag{34}$$

The second law efficiency of the cycle can also be expressed as the ratio of the first law efficiency and the Carnot efficiency. Accordingly, we can write:

$$\eta_{II,RRCBC} = \eta_{I,RRCBC} / \eta_{Carnot,RRCBC} \tag{35}$$

where, the Carnot efficiency can be written as [46]:

$$\eta_{Carnot,RRCBC} = 1 - T_0 / T_{g,in} \tag{36}$$

ACCEPTED MANUSCRIPT

If an equal value is obtained for each of the above expressions for the second law efficiency, it should indicate a proper exergy balance. The first law efficiency of WHRS can be expressed as the ratio of the net power produced by the system to the total energy contained in the incoming gas stream in the HRHE, while the second law efficiency of WHRS can be defined as the ratio of the net power produced by the system to the total availability of the incoming gas stream in the HRHE, as given below [44]:

$$\eta_{I,WHRS} = \dot{W}_{net} / \left(\dot{m}_g C_{p,g,in} T_{g,in} \right) \tag{37}$$

$$\eta_{II,WHRS} = \dot{W}_{net} / \left(\dot{m}_g e_{g,in,RRCBC} \right) \tag{38}$$

where $T_{g,in}$ and $C_{p,g,in}$ are the gas inlet temperature and the isobaric specific heat of the gas, respectively, $e_{g,in}$ represents the specific exergy of gas at the HRHE inlet, and the denominator on the right hand side of the above equation represents the total availability of gas at the HRHE inlet.

2.4 Exhaust gas composition analysis

The details of the different gas composition scenarios are presented in Table 2. The isobaric molar specific heats of the component ideal gases contained in the exhaust gas mixture are determined using the following relations for the non-NOx gases [45] and the NOx gases [46], respectively:

$$\bar{C}_{pi} = \frac{R}{M} \left(\alpha + \beta T + \gamma T^2 + \delta T^3 + \varepsilon T^4 \right) \tag{39}$$

$$\bar{C}_{pi} = \left(a + bT + cT^2 + dT^3 \right) \tag{40}$$

where $\bar{C}_{p,i}$ is the isobaric molar specific heat of the ith component gas; $\alpha, \beta, \gamma, \delta, \varepsilon$ and a, b, c, d are the gas constants used for calculation of isobaric specific heat of the non-NOx gases [45] and the NOx gases [46] respectively, T is the absolute temperature of the gas mixture, R is the universal gas constant (8.31451 kJ.kmol^{-1}K^{-1}); M is the molar mass (kg.kmol^{-1}) of the respective ideal gas; Eq.39 is valid for temperatures from 300 to 1000 K and Eq.40 is valid for temperatures from 273 to 1500 K. The specific heat, $C_p(T)$ and entropy change, $(s_2 - s_1)$ of the exhaust gas mixture, can thus be expressed as [45]:

$$C_p(T) = \sum_{i=1}^{n} y_i C_{p_i}(T) \tag{41}$$

ACCEPTED MANUSCRIPT

$$s_2 - s_1 = \sum_{i=1}^{n} y_i (s_2 - s_1)_i \quad (42)$$

where y_i is the mass fraction (%), $C_{pi}(T)$ is the isobaric specific heat, and $(s_2 - s_1)_i$ is the entropy change, of each ideal gas component of the gas mixture.

3. Simulation and validation of model

The thermodynamic analyses of RRCBC and WHRS has been carried out by simulating an in-house computer code developed in Engineering Equation Solver (EES) software [48]. For thermodynamic properties of CO_2, the EES in-built library is used, wherein the values of specific enthalpy and entropy are referred to values of 0 at 298.15 K and 101.325 kPa in the ideal gas state. It provides high accuracy thermodynamic properties for CO_2 using the fundamental equation of state developed in the literature [49]. The models search for the state points for all the components of the cycle using the input parameters as per the set of standard operating & design condition presented in Table 3. Unless otherwise specified, these conditions will be the input parameters for performance analyses of the bottoming cycle and WHRS. The energy and exergy balance checks have been incorporated in the model to verify proper energy and exergy balance. For the HRHE and the precooler, the temperature differences are defined in the model and initially set to 22 ^{0}C and 15 ^{0}C respectively, in order to avoid the pinch effect during heat transfer in these components. For regenerators (HTR and LTR), the temperature differences at the hot and the cold ends have been defined, in order to assess the pinch effect in these components. The care was taken to ensure that these temperature differences are always greater than or equal to 20 ^{0}C, in order to avoid the pinch effect during heat transfer. However, the exhaust gas outlet temperature in HRHE was kept above its saturation temperature of 549.15 K (276 ^{0}C). The dead or the reference state, represented by temperature (T_o) and pressure (P_o) has been assumed as 308.15 K (35 ^{0}C) and 1.0 bar, respectively.

The optimization variable considered during simulation is the cycle pressure ratio (r_p) which is defined as the ratio of maximum cycle pressure and the minimum cycle pressure i.e., P_{max}/P_{min}. The optimization has been carried by determining the optimum pressure ratio ($r_{p,\,opt}$) corresponding to the maximum exergetic efficiency of the cycle. For simulating the model, the cycle pressure ratio was varied in steps of 0.01 bar with an upper bound of 40 bar and lower bound of 20 bar, while the maximum cycle pressure was fixed at 200 bar, since the higher working pressure implies more complexity and cost. The mathematical

ACCEPTED MANUSCRIPT

model developed for RRCBC has been verified with the published literature [9] and prototype design data [13]. A summary of comparison of results is placed in Table 4. It is found that the model results are in good agreement with the literature, though a slight deviation in values may be attributable to the heat exchangers pressure drop (130 kPa) considered in the literature [9] against nil pressure drop assumed in the present results. The WHRS model is also verified with the published literature [44] and the results are found to be closely matching with the literature.

4. Results and Discussions

The mathematical model developed for regenerative recompression Brayton cycle (RRCBC) has been presented and studied for a typical set of operating parameters and also validated with the already published work given in the literature. The set of input parameters assumed are based on those already given in the literature and are shown in Table 3 and the discussion of results is given as below:

4.1 Effects of critical point on specific heat

Figure 3(a) illustrates the variation of specific heat capacities of CO_2 with the temperature and the pressure, in the vicinity of its critical point. As can be clearly seen, the specific heat capacities are found to be significantly varying with the temperature and also with pressure, especially, near the critical point temperature and pressure. The specific heat capacity is found to be reaching a maximum value at the critical point and then decreasing when the temperature and pressure move away from the critical point. These sharp variations in the specific heat capacities of CO_2 are due to its real gas behaviour in the vicinity of the Andrews curve, which has a significant influence on its transport properties and on heat transfer coefficients [4, 9]. Therefore, even a slight change in the temperature and the pressure in a thermodynamic power conversion cycle employing CO_2 will result in significant changes in the thermo-physical and transport properties of CO_2. This in turn will affect the heat transfer characteristics and the cycle performance. Therefore, although the supercritical CO_2 cycles have distinct advantages due to these significant changes in the properties near the critical point region, particular care is also required in these cycles to ensure that the thermal processes near the critical point e.g., compression are performed at

ACCEPTED MANUSCRIPT

the optimum conditions [9] which implies that the cycle pressure ratio is an important variable that must be studied in detail for the supercritical CO_2 cycles.

4.2 Effects of cycle pressure ratio (r_p)

As observed above, the cycle pressure ratio is an important variable that needs a careful study for the supercritical CO_2 cycles. Therefore, the influence of cycle pressure ratio (r_p) on the energetic and the exergetic efficiency of the proposed supercritical CO_2 power cycle has been evaluated and illustrated in Fig. 3(b). It is seen that the cycle is quiet sensitive to the pressure ratio and that even a small departure from the optimum pressure ratio can cause a significant reduction in cycle performance. This trend is on account of two factors. Firstly, for a fixed heat exchanger volume as the pressure ratio is varied, the heat exchangers (including HTR and LTR) effectiveness (and thus cycle efficiency) is affected because of the variations in the working fluid density which affects the velocity and pressure drops. Thus, at high pressure ratios, the density of the working fluid at the high pressure side of heat exchangers becomes lower resulting in higher velocity and higher pressure drops. The steep increase in pressure drops with the pressure ratio at high pressure side of the heat exchangers (including LTR and HTR) contributes significantly to the deterioration of the cycle efficiency at high pressure ratios. For a fixed heat exchanger volume, at low pressure ratios, it is the HTR that is responsible for the significant reduction in cycle efficiency and at high pressure ratios; the LTR also contributes albeit lower than HTR's contribution. Secondly, the heat available for regeneration also decreases as the pressure ratio increases which also contributes in reduction in the overall cycle efficiency. Similar results have already been reported elsewhere, in the literature [9].

4.3 Exergy destruction analysis

The temperature-entropy and the temperature-exergy diagrams for the cycle are shown in Figs. 2(a) and 2(b), respectively. Table 5 presents the thermodynamic properties at defined state points, as shown in Fig. 1, for RRCBC at standard operating and design condition and cycle pressure ratio=2.5. On the other hand, the exergy balance of RRCBC at the optimum cycle pressure ratio of 2.47, without pressure drop in the heat exchangers, is presented in Table 6. The temperature-exergy diagram (Fig. 2(b)) shows that the exergy variation in the precooler and HRHE are significant as compared to other components. The exergy variation in the precooler is more pronounced as compared to other components,

ACCEPTED MANUSCRIPT

while it is found to be the reversed in the case of recompression compressor, which is due to the higher temperature difference in the precooler.

Similarly, it is found that the irreversibility in the heat exchangers is significantly more as compared to that of the turbo-machineries. On the other hand, among the regenerators, the exergy destruction in the HTR is more pronounced as compared to that of the LTR. In contrast, the HRHE, precooler and regenerators altogether contribute about 80% of total exergy destruction within the cycle, which is due to the higher heat exchange temperature difference among these components. Therefore, from the exergetic point of view, these subsystems are the most critical components of a supercritical CO_2 based waste heat recovery Brayton cycle for power generation in the shipboard application. Similarly, the exergetic loss in the turbine is found to be more than that of the compressors together, which is due to the fact that the temperature difference in the former is much higher than the latter leading to the results as mentioned in Table 6.

4.4 Effects of compressor inlet temperature (T_5)

Figure 4(a) depicts the dependence of the optimum cycle pressure ratio and cycle efficiency on the compressor inlet temperature (T_5). The results also show that the cycle efficiency decreases almost linearly with the increase in the compressor inlet temperature. This is because of the fact that the temperature difference between the turbine inlet and the compressor (viz. the temperature lift) decreases linearly with the increase in the compressor inlet temperature.

It is also seen from Fig. 4(a) that the optimum pressure ratio sharply decreases with the increase in the compressor inlet temperature up to 40 ^{0}C and then slightly increases throughout. This is due to the fact that the critical point of the working fluid is around 31 ^{0}C and the specific heat, which plays a very important role in the efficiency and heat transfer process, varies sharply in this temperature range, mentioned above, leading to the result obtained here. In other words, if the cycle is designed for a certain compressor inlet temperature and the operation is performed at different compressor inlet temperature, it will result in a significant decrease in the cycle efficiency, as the cycle will be operating away from its optimum pressure ratio [9]. Thus, the increase in the minimum cycle temperature from 32 to 50 ^{0}C leads to reduction in the cycle efficiency by 11%, showing the above mentioned effect of sharp variation in the specific heat.

ACCEPTED MANUSCRIPT

The effect of the compressor inlet temperature on the irreversibility of different components is shown Fig. 4(b) for a typical set of operating parameters, mentioned above. It is clear from Fig. 4(b) that the irreversibility of the precooler increases sharply while for the main compressor it increases, slightly. This particular change in the result is due to the sudden changes in the specific heat of the working fluid, as discussed above. On the other hand, the irreversibility associated with five other components such as, HRHE, LTR, turbine, and recompressor decreases slightly. The changes in the latter five components is very less as compared the former two components, mentioned above. Whereas, for the high temperature regenerator, the irreversibility first increases up to 38 ^{0}C and then decreases at a faster pace. This particular result is very likely due to the sudden changes in the specific heat of the working fluid in the vicinity of this temperature range, as already explained in the section 4.1, above.

The physical significance of the above results can be explained in different ways. For example, in the case of precooler, with the increase in temperature, the temperature difference between the hot stream (CO_2) and the cold stream (cooling water) increases significantly which leads to increase in irreversibility. Further, it is observed that, up to a certain temperature, the precooler irreversibility increases sharply and then modestly which is basically due to the sharp variation in the specific heat of the working fluid in the vicinity. However, when sufficiently far away from the critical point, the degradation of the specific heat of SC-CO_2 affects the heat transfer in precooler and thus the irreversibility rises modestly in this region. The irreversibility of the heat source (HRHE) decreases as the average heat exchange temperature difference (between the gas and CO_2 inlet to HRHE) decreases as the temperature of CO_2 at HRHE inlet increases with rise in main compressor inlet temperature. In the present analysis, the irreversibility in precooler is observed to be higher than that of HRHE(after about 40 ^{0}C) as the precooler pumping power is high due to high mass flow rate of cooling water and high heat transfer in the precooler owing to high average heat exchange temperature difference between the CO_2 and the cooling water.

4.5 Effects of turbine inlet temperature (T_1)

Figure 5(a) illustrates the effect of turbine inlet temperature (T_1) on the optimum cycle pressure ratio and corresponding efficiencies. As the maximum cycle temperature increases, the energetic as well as exergetic efficiency increases almost linearly. This is also

ACCEPTED MANUSCRIPT

obvious because of the fact that the increase in turbine inlet temperature increases the opportunity to do useful work. In other words, the potential of doing useful work per unit time per cycle increases leading to the results obtained and shown in Fig. 5(a). Further from the second law of thermodynamics, the temperature-lift of the working fluid, by each degree enhances the quality of energy and hence, the potential to do useful work viz. the availability of the stream across the turbine leading to higher exergetic efficiency of the system.

Similarly, the effect of the turbine inlet temperature on the optimum pressure ratio is negligible as compared to that of the compressor inlet temperature, which is again due to the fact that the specific heat of the working fluid (CO_2) in the vicinity of the turbine is nearly invariant, as it is located far away from the critical point, unlike the compressor inlet temperature. Further, the increase in maximum operating temperature by 100 ^{0}C, causes an increase in cycle efficiency by almost 10%, indicating that the effect due to variation in the maximum cycle temperature is quite significant, as discussed earlier, based on the quality (availability or exergy) of the working fluid at that particular state point.

Further, the effect of the turbine inlet temperature on the component irreversibility is illustrated in Fig. 5(b). It is clear from Fig. 5(b) that, with an increase in the turbine inlet temperature, the irreversibility of the HRHE decreases sharply while for the HTR, it decreases slightly since the specific heat capacity values of CO_2 at all state points are nearly invariant with the increase in turbine inlet temperature, as they are located far away from the critical point. Therefore, all the components except the precooler and HRHE, have an almost negligible effect on the irreversibility. On the other hand, the irreversibility increases slightly for precooler, while for other components such as turbine, compressors and the LTR, the variation is almost negligible. Further, for the case of the HRHE, the irreversibility shows a sharp decreasing trend, which is due to the fact that an increase in outflow enthalpy results in an enhanced heat transfer in the HRHE, as also explained earlier, leading to the results shown in Fig. 5(b).

4.6 Effects of pressure drop on cycle performance

The effect of different pressure drop scenarios ($\delta\bar{P}_{ij}$=0, 1% and 3%) on the performance parameters at the respective optimal pressure ratios is shown in Table 7. It is found from the table that 3% increase in the relative pressure drop fraction, results in

ACCEPTED MANUSCRIPT

about 7% increase in the optimum cycle pressure ratio, which is quite significant. It is also found from the table that the effect of the pressure drop is more pronounced for the net power output, while it was found to be least pronounced for the mass flow rate of the working fluid. Whereas, the effects of the pressure drop on the energetic and exergetic efficiencies of the two different systems viz. RRCBC and WHRS is found to be significant. In the case of RRCBC, the energy and exergy efficiency vary by 7% and 12%, while in the case of WHRS they are found to be 2% and 8%, respectively. However, a small pressure drop of 3% results in decrease of total power output by 25% which is quite significant and leads to higher power penalty of the entire system.

Further, the effects of the pressure drop across the different heat exchangers on the exergetic efficiency of RRCBC is also graphically shown in Fig. 6. It is seen from Fig. 6 that the effects of the increase in pressure drop from 0 to 3%, in different heat exchangers viz. HRHE, HTR, LTR and precooler, on the cycle exergetic efficiency is found to be in the range of 3.3% to 2.2%, which, though not very significant, however, cannot be neglected as well. It is clearly seen from Fig. 6 that the exergetic efficiency of the cycle decreases linearly with the increase in pressure drop across the different heat exchangers. The results also show that the effect of pressure drop within the HRHE and the HTR is more significant followed by the LTR, while, the effect due to precooler pressure drop is less pronounced on the degradation of the exergetic efficiency. This can also be explained in different ways, for example, in the precooler, the mass flow rated is bifurcated so the net changes in the availability which is mass flow rate times the exergy is less and hence we get the results as obtained above.

4.7 Effects of gas compositions

In order to analyse the influence of the exhaust gas compositions and the gas inlet temperature on the performance of the waste heat recovery system, the heat input to the topping cycle was varied from 58.7 to 42.2 MW in steps of below 1.0 MW, keeping the gas mass flow rate and turbine inlet temperature of the topping cycle as constants. The three gas composition scenarios namely, the Air-Standard [44], Gas-I (actual gas with 15.63% oxygen mass concentration) and Gas-II (actual gas with 7.81% oxygen mass concentration) were considered for the gas composition analysis. Figure 7 illustrates the effects of the

ACCEPTED MANUSCRIPT

three different gas compositions and the gas inlet temperature (T_{gin}) on the energetic and exergetic efficiency of the waste heat recovery system (WHRS). The results clearly show that the analysis based on air standard approach underestimates the first law efficiency, while it significantly overestimates the second law efficiency and therefore, the actual exhaust gas composition should be used to accurately predict the first and the second law efficiencies of a WHRS. Thus, it would be better if there is some online reading available, using some smart sensors, for the actual exhaust gas compositions, for the better optimization of the heat recovery, load on the topping cycle and hence, the potential and actual output of the bottoming cycle, for the best operating parameter and the maximum performance of the entire system for shipboard applications.

5. Conclusions

In the current work, the modelling and analysis of waste-heat-recovery-system (WHRS) using supercritical carbon dioxide (SC-CO_2) regenerative recompression Brayton cycle (RRCBC) has been presented for shipboard power applications based on energy and exergy analyses. The influence of key operating parameters such as, gas composition, turbine and compressor inlet temperatures, pressure drop irreversibility, pressure ratio etc., on the overall performance of the system including the exergy destruction rate has been investigated. It is found that the exergy destruction in the heat exchangers is significantly higher which is about four times higher as compared to that in the turbomachinery. From the exergy balance, it is found that the precooler, HRHE and the regenerators are the most important components from the exergetic performance point of view. Further, it is found that there is an optimum pressure ratio at which the cycle attains the maximum efficiency for a typical set of operating parameter. Again, the variation in any of the operating parameters, will result in the significant changes not only in the optimum pressure ratio but also the corresponding thermal and exergetic efficiency of the cycle.

In other words, the cycle is found to be quiet sensitive to the pressure ratio and that even a small departure from the optimum pressure ratio can cause a significant reduction in cycle performance. The compressor inlet temperature has been found to have a more effect, not only on the optimum cycle pressure ratio but also on the cycle performance, than that of the turbine inlet temperature. It is clearly seen from study that the exergetic efficiency of the cycle decreases linearly with the increase in pressure drop across the

ACCEPTED MANUSCRIPT

different heat exchangers. The results also show that the effect of pressure drop within the HRHE and the HTR is more significant followed by the LTR, while, the effect due to precooler pressure drop is less pronounced on the degradation of the exergetic efficiency. In other words, the existence of the pressure drop leads to the drop in the availability (exergy) of the stream thereby affecting the performance significantly.

Finally, the energy and exergy based performance evaluation presented in this cycle model is unique and different from those already done by earlier workers as available in the literature, for a typical set of operating parameter. In other words, the results obtained by earlier worker can be easily drawn from the present cycle model as a special case and hence, this cycle model is more general and different from those of the already published work.

Further, in order to have the experimental validity of the proposed work, a prototype will be required to validate the design concepts experimentally and to get realistic testing data for transient analysis and experimental studies for shipboard gas turbine and the work in this direction is on-going. Moreover, the component sizing, thermo-economic optimisation, dimensional feasibility study of the proposed system, towards its shipboard installation and its integration with the ship's systems, will also be required to be carried out in future.

Nomenclature

a, b, c, d	gas constants used for calculation of isobaric specific heat of (ideal) NO_x gas [46]	TEG	thermo-electric-generator
ABTR	advanced burner test reactor	$\dot{W}$	power, kW
Ar	argon	WG	water-gauge (pressure) (inch)
BWR	back-work-ratio	WHRS	waste-heat-recovery-system
C_p	isobaric specific heat (kJ.kg^{-1}.K^{-1})	x	recompression-mass-fraction
$\overline{C}_p$	isobaric molar specific heat (kJ.kmol^{-1}.K^{-1})	y	mass fraction (%)
			Subscripts and superscripts
CO	carbon mono oxide		
CO_2	carbon dioxide (R744)		
CSP	concentrated solar power	*Carnot*	ideal or Carnot efficiency
e	specific exergy(kJ.kg^{-1})	*CC*	combined cycle
$\dot{E}$	exergy rate (kW)	CO_2	carbon dioxide (working fluid)
GWP	global warming potential	*g*	exhaust gas
h	specific enthalpy (kJ.kg^{-1})	*HRHE*	heat-recovery-heat-exchanger
HRHE	heat-recovery-heat-exchanger	*HTR*	high-temperature-regenerator

ACCEPTED MANUSCRIPT

HRSG	heat-recovery-steam-generator	I	first-law or energetic or thermal efficiency
i_j	non-dimensional irreversibility fraction for the j^{th} component	II	second-law or exergetic efficiency
$\dot{I}$	Irreversibility or exergy destruction rate (kW)	$isen$	isentropic
LSHSD	low sulphur high speed diesel (marine diesel fuel)	i	states (1 - 8) of the cycle or gas component
$\dot{m}$	mass flow rate ($kg.s^{-1}$)	in	inlet of control volume
M	molar mass ($kg.kmol^{-1}$)	j	component
NO	nitrogen oxide	LTR	low-temperature-regenerator
NO_2	nitrogen di oxide	MC	main-compressor
NOx	mixture of NO and NO_2 etc.	net	net value
ODP	ozone depletion potential	opt	optimum value
O_2	oxygen	out	outlet of control volume
p	pressure, MPa	RC	re-compression compressor or recompressor
PPTD	pinch-point-temperature-difference (^{0}C or K)	$RRCBC$	regenerative re-compression Brayton (power) cycle
$\dot{Q}$	heat transfer rate (kW)	s	ideal or isentropic state
r_p	pressure ratio (p_{max}/p_{min})	w	cooling water
$r_{p,opt}$	optimum pressure ratio	$WHRS$	waste-heat-recovery-system
R	universal gas constant ($kJ.kmol^{-1}.K^{-1}$)	$1,2,..8$	cycle states [Figs. 1, 2(a) and 2(b)]
RACER	rankine-cycle-exhaust-recovery-system	0	dead state
RH	relative humidity (%)		*Greek letters*
s	specific entropy ($kJ.kg^{-1}.K^{-1}$)		
SC	super-critical	$\alpha, \beta, \gamma, \delta$	gas constants used for calculation of isobaric specific heat of (ideal) non-NOx gas [45]
SFR	sodium-cooled-fast-reactor	ε	effectiveness (%)
SO_2	sulphur dioxide	η	efficiency (%)
SSSF	steady-state-steady-flow	ΔP	pressure drop across a stream($i-j$)
T	Temperature (K)	$\delta \bar{P}_{ij}$	non-dimensional relative pressure drop fraction

Acknowledgement

The authors are thankful to Dr. S.K. Tyagi for his valuable suggestions during various discussions from time to time.

References

ACCEPTED MANUSCRIPT

1. L. Pierobon, A. Benato, E. Scolari, F. Haglind, A. Stoppato, Waste heat recovery technologies for offshore platforms, Applied Energy 136 (2014) 228-41.
2. E.G. Feher, The supercritical thermodynamic power cycle, presented to the IECEC, Florida, Douglas paper no. 4348 (1967).
3. E.G. Feher, The supercritical thermodynamic power cycle, Int. J. Energy Conversion 8 (1968) 85-90.
4. G. Angelino, Real gas effects in carbon dioxide cycles, ASME, Gas Turbine Conf. and Products Show V001T01A071 (1969).
5. O.V. Combs, An investigation of the supercritical CO_2 cycle (Feher cycle) for shipboard application, Doctoral dissertation, Massachusetts Institute of Technology (1977).
6. H. D. Marron, Gas turbine waste heat recovery propulsion for U.S. Navy surface combatants, Naval Engineers J. 93(5) (1981) 65–75.
7. V. Dostal, M.J. Driscoll, P. Hejzlar, N.E. Todreas, A supercritical CO_2 gas turbine power cycle for next-generation nuclear reactors, ASME, Proceedings of International Conference of Nuclear Engineering (ICONE10-22192), 1st Jan 2002, Arlington, Virginia, (2002) 567-74.
8. V. Dostal, M.J. Driscoll, P. Hejzlar, A supercritical carbon dioxide cycle for next generation nuclear reactors. Massachusetts Institute of Technology, Dept. of Nuclear Eng., Cambridge, Paper No. MIT-ANP-TR-100, MA (2004).
9. V. Dostal, A supercritical carbon dioxide cycle for next generation nuclear reactors, Ph.D. thesis, MIT, USA (2004).
10. L. Moroz, M. Burlaka, O. Rudenko, C. Joly, Evaluation of gas turbine exhaust heat recovery utilizing composite supercritical CO_2 cycle, Proceedings of International Gas Turbine Congress, Tokyo, Japan, 15th-20th Nov 2015, ISBN978-4-89111-008-6 (2015).
11. Di Bella, F. A, Gas turbine engine exhaust waste heat recovery Navy shipboard module development, International Supercritical CO_2 Power Cycle Symposium, Boulder, Colorado, 24th May 2011.
12. Di Bella, F. A, Gas turbine engine exhaust waste heat recovery using supercritical CO_2 Brayton cycle with thermoelectric generator technology, ASME, 9th International Conference of Energy Sustainability, 28th Jun-2nd Jul 2015, V001T04A003 (2015).
13. S. Ishiyama, Y. Muto, Y. Kato, S. Nishio, T. Hayashi, Y. Nomoto, Study of steam, helium and supercritical CO_2 turbine power generations in prototype fusion power reactor, Progress in Nuclear Energy 50(2) (2008) 325-32.
14. J. Sarkar, Second law analysis of supercritical CO_2 recompression Brayton cycle, Energy 34(9) (2009) 1172-8.
15. A. Moisseytsev, J.J. Sienicki, Investigation of alternative layouts for the supercritical carbon dioxide Brayton cycle for a sodium-cooled fast reactor, Nuclear Engineering Design 239(7) (2009) 1362-71.
16. J.J. Dyreby, S.A. Klein, G.F. Nellis, D.T. Reindl, Modelling off-design and part-load performance of supercritical carbon dioxide power cycles, InASME Turbo Expo 2013, Turbomachinery Technical Conference and Exposition, 3rd Jun 2013, V008T34A014 (2013)

ACCEPTED MANUSCRIPT

17. J. Dyreby, S. Klein, G. Nellis, D. Reindl, Design considerations for supercritical carbon dioxide Brayton cycles with recompression. Journal of Engineering for Gas Turbines and Power 136 (10) (2014) 101701.
18. H.S. Pham, N. Alpy, J.H. Ferrasse, O. Boutin, J. Quenaut, M. Tothill, D. Haubensack, M. Saez, Mapping of the thermodynamic performance of the supercritical CO_2 cycle and optimization for a small modular reactor and a sodium-cooled fast reactor, Energy 87 (2015) 412-424.
19. H. Zhao, Q. Deng, W. Huang, D. Wang, Z. Feng, Thermodynamic and economic analysis and multi-objective optimization of supercritical CO_2 Brayton cycles, Journal of Engineering for Gas Turbines and Power, 138(8) (2016) 081602.
20. X.R. Zhang, H. Yamaguchi, D. Uneno, K. Fujima, M. Enomoto, N. Sawada, Analysis of a novel solar energy-powered Rankine cycle for combined power and heat generation using supercritical carbon dioxide, Renewable Energy 31(12) (2006) 1839-54.
21. P. Garg, P. Kumar, K. Srinivasan, Supercritical carbon dioxide Brayton cycle for concentrated solar power, The Journal of supercritical Fluids 76(2013) 54-60.
22. C.S. Turchi, Z. Ma, T.W. Neises, M.J. Wagner, Thermodynamic study of advanced supercritical carbon dioxide power cycles for concentrating solar power systems, Journal of Solar Energy Engineering 135(4) (2013) 041007.
23. T.W. Neises, C.S. Turchi, A comparison of supercritical carbon dioxide power cycle configurations with an emphasis on CSP applications, Energy Procedia 49 (2014) 1187-96.
24. R.V. Padilla, Y.C. Too, R. Benito, W. Stein, Exergetic analysis of supercritical CO_2 Brayton cycles integrated with solar central receivers, Applied Energy 148(2015) 348-65.
25. Y. Chen, P. Lundqvist, P. Platell, Theoretical research of carbon dioxide power cycle application in automobile industry to reduce vehicle's fuel consumption, Applied Thermal Engineering 25(14) (2005) 2041-53.
26. D. Sanchez, J.M.M. de. Escalona, R. Chacartegui, A. Munoz, T. Sanchez, A comparison between molten carbonate fuel cells based hybrid systems using air and supercritical carbon dioxide Brayton cycles with state of the art technology, Journal of Power Sources 196(2011)4347-54.
27. E.M. Clementoni, T.L. Cox, M.A. King, Off-nominal component performance in a supercritical carbon dioxide Brayton cycle, Journal of Engineering for Gas Turbines and Power 138 (2016) 011703.
28. F. Lambruschini, E. Liese, S.E. Zitney, A. Traverso, Dynamic Model of a 10 MW supercritical CO_2 recompression Brayton cycle, InASME Turbo Expo 2016, Turbomachinery Technical Conference and Exposition, GT2016, 13^{th}-17^{th} Jun 2016 V009T36A004(2016).
29. Y. Ahn, S.J. Bae, M. Kim, S.K. Cho, S. Baik, J. Lee, J.E. Cha, Review of supercritical CO_2 power cycle technology and current status of research and development, Nuclear Engineering and Technology 31(47) (2015) 647-61.

ACCEPTED MANUSCRIPT

30. V. Ganapathy, Heat-recovery steam generators: understand the basics, Chemical Engineering Progress (1996) 32–45.
31. T.H. Ko, K. Ting, Optimal Reynolds number for the fully developed laminar forced convection in a helical coiled tube, Energy 31(12) (2006) 2142-52.
32. G. Xie, Y. Song, M. Asadi, G. Lorenzini, Optimization of pin-fins for a heat exchanger by entropy generation minimization and constructal law, Journal of Heat Transfer 137(6) (2015) 061901.
33. H. Najafi, B. Najafi, P. Hoseinpoori, Energy and cost optimization of a plate and fin heat exchanger using genetic algorithm, Applied Thermal Engineering 31(10) (2011) 1839-47.
34. M.R. Hajmohammadi, E. Shirani, M.R. Salimpour, A. Campo, Constructal placement of unequal heat sources on a plate cooled by laminar forced convection, International Journal of Thermal Sciences 60 (2012)13-22.
35. M.R. Hajmohammadi, S.S. Nourazar, A. Campo, S. Poozesh, Optimal discrete distribution of heat flux elements for in-tube laminar forced convection, International Journal of Heat and Fluid Flow 40 (2013) 89-96.
36. M.R. Hajmohammadi, M. Rahmani, A. Campo, O.J. Shariatzadeh, Optimal design of unequal heat flux elements for optimized heat transfer inside a rectangular duct, Energy 68 (2014) 609-16.
37. M.R. Hajmohammadi, G. Lorenzini, O.J. Shariatzadeh, C. Biserni, Evolution in the design of V-shaped highly conductive pathways embedded in a heat-generating piece, Journal of Heat Transfer 137(6) (2015) 061001.
38. M.R. Hajmohammadi, O.J. Shariatzadeh, M. Moulod, S.S. Nourazar, Phi and Psi shaped conductive routes for improved cooling in a heat generating piece, International Journal of Thermal Sciences 77 (2014) 66-74.
39. S.C. Kaushik, S.K. Tyagi, Finite time thermodynamic analysis of an irreversible regenerative closed cycle Brayton heat engine, Int. J. of Solar Energy 22 (3-4) (2002) 141-51.
40. S.C. Kaushik, S.K. Tyagi, M.K. Singhal, Parametric study of an irreversible regenerative Brayton cycle with isothermal heat addition, Energy Conversion and Management 44 (2003) 2013-25.
41. S.K. Tyagi, G.M. Chen, Q. Wang, and S.C. Kaushik, Thermodynamic analysis and parametric study of an irreversible regenerative-intercooled-reheat Brayton cycle, International Journal of Thermal Sciences 45 (2006) 829-40.
42. T.J. Kotas, The exergy method of thermal plant analysis, Elsevier (2013).
43. A. Bejan, Fundamentals of exergy analysis, entropy generation minimization, and the generation of flow architecture, International Journal of Energy Research 26(7) (2002) 545-65 (DOI: 10.1002/er.804)
44. C.J. Butcher, B.V. Reddy, Second law analysis of a waste heat recovery based power generation system, International Journal of Heat and Mass Transfer 50(11) (2007) 2355-63.

ACCEPTED MANUSCRIPT

45. M.J. Moran, H.N. Shapiro, D.D. Boettner, M.B. Bailey, Fundamentals of engineering thermodynamics, John Wiley & Sons (2010).
46. Y.A. Cengel, M.A. Boles, Thermodynamics – an engineering approach, Tata McGraw Hill, New Delhi (2002).
47. S.A. Klein, EES, Engineering Equation Solver, version 8.158, F-Chart Software, Middleton, WI (2008).
48. R. Span, W. Wagner, A new equation of state for carbon dioxide covering the fluid region from the triple point temperature to 1100 K at pressures up to 800 MPa, Journal of Physical and Chemical Reference Data 25(6) (1996) 1509-96.
49. Internet website, http://www.geaviation.com/marine/engines/military/lm2500 regarding specifications of shipboard aeroderivative marine gas turbine engine (Last accessed on 07th Jan 2017).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

List of Figures

S. No.	Figure Description
Fig. 1.	Proposed shipboard WHRS and RRCBC (Thermodynamic properties at defined state points shown in Table 5).
Fig. 2.	(a) Temperature- entropy diagram; (b) Temperature-specific exergy diagram, of RRCBC.
Fig. 3.	(a) Variation of specific heat of CO_2 in the vicinity of the critical point; (b) Effect of cycle pressure ratio on the cycle performance.
Fig. 4.	Effect of minimum cycle (compressor inlet) temperature (T_5) on (a) cycle performance and optimum cycle pressure ratio ($r_{p,\,opt}$); (b) component irreversibility.
Fig. 5.	Effect of maximum cycle (turbine inlet) temperature (T_1) on (a) cycle efficiency and optimum cycle pressure ratio ($r_{p,\,opt}$); (b) component irreversibility.
Fig. 6.	Effect of component pressure drops on cycle exergetic efficiency.
Fig. 7.	Effect of gas compositions and gas inlet temperature on (a) the energetic efficiency (b) the exergetic efficiency, of WHRS.

List of Tables

S. No.	Table Description
Table 1.	Broad specifications of the shipboard marine gas turbine Ref. [49].
Table 2.	Details of exhaust gas composition scenarios.
Table 3.	Standard operating and design condition.
Table 4.	Comparison of results with the published literature.
Table 5.	Thermodynamic properties at defined state points for supercritical CO_2 regenerative recompression Brayton cycle, at standard operating and design condition.
Table 6.	Exergy balance for RRCBC at standard operating and design condition, without pressure drop, and optimum cycle pressure ratio=2.47.
Table 7.	Optimized performance results at standard operating and design condition at different pressure drop fractions.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

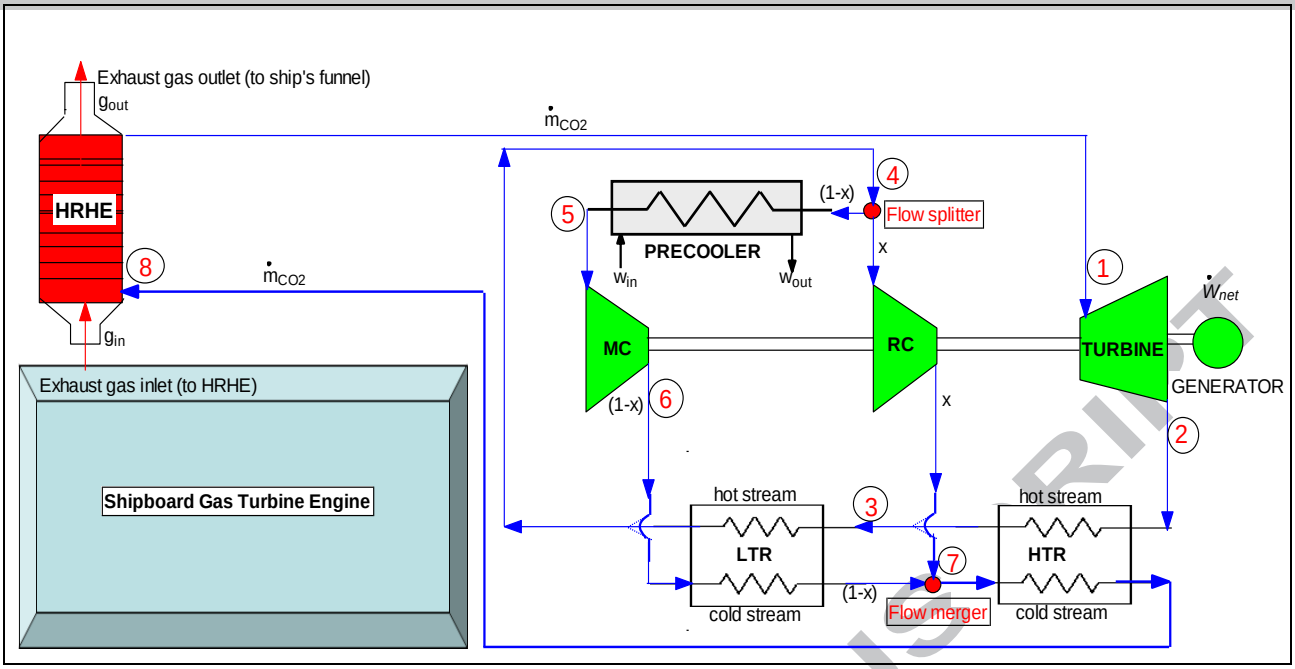


Fig. 1. Proposed shipboard WHRS with RRCBC (Thermodynamic properties at defined state points shown in Table 5).

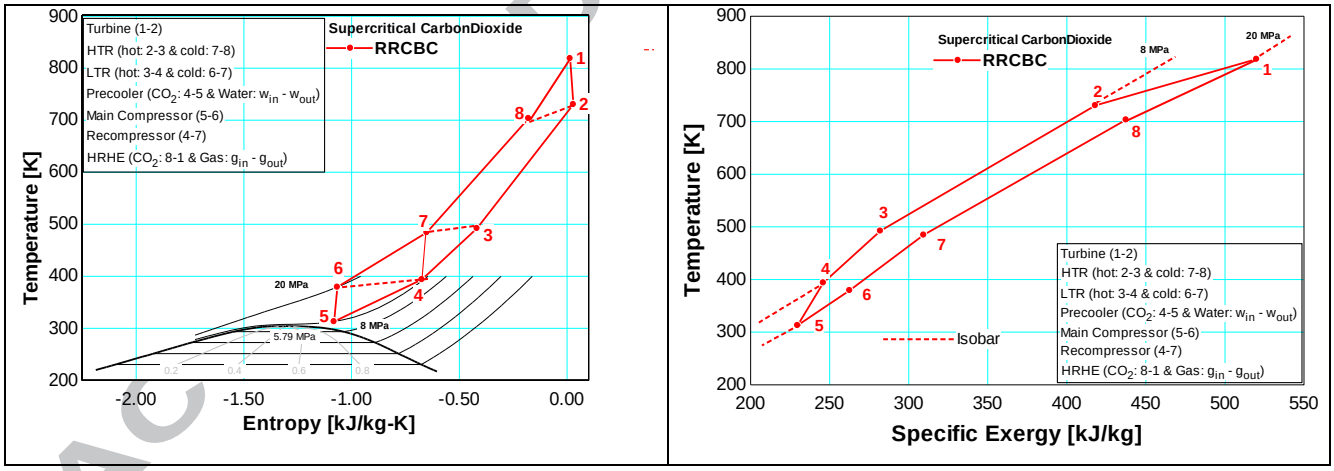


Fig. 2. (a) Temperature- entropy diagram; **(b)** Temperature-specific exergy diagram, of RRCBC.

ACCEPTED MANUSCRIPT

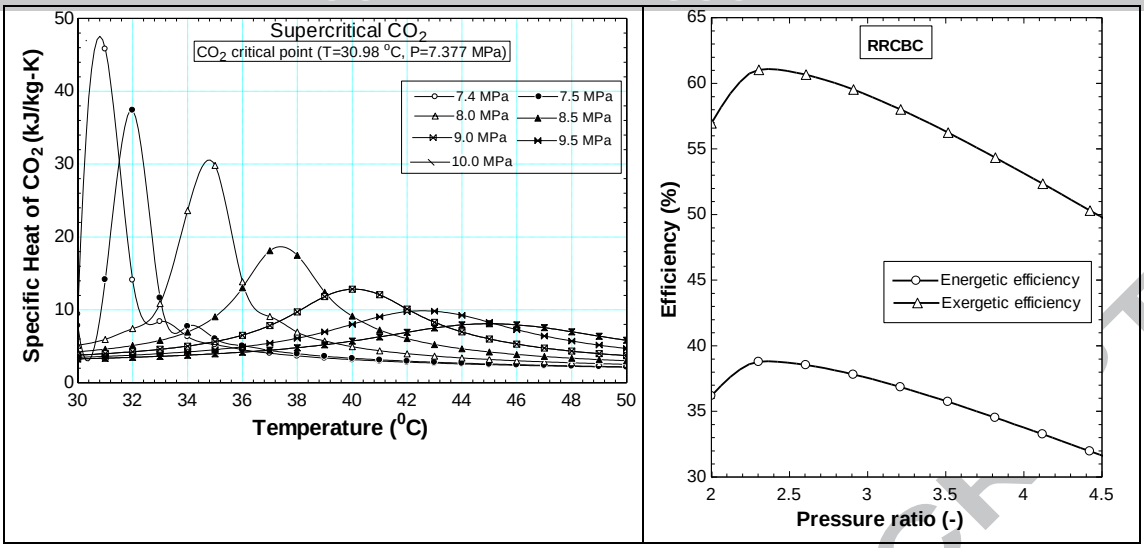


Fig. 3. **(a)** Variation of specific heat of CO_2 in the vicinity of the critical point; **(b)** Effects of cycle pressure ratio on the cycle performance.

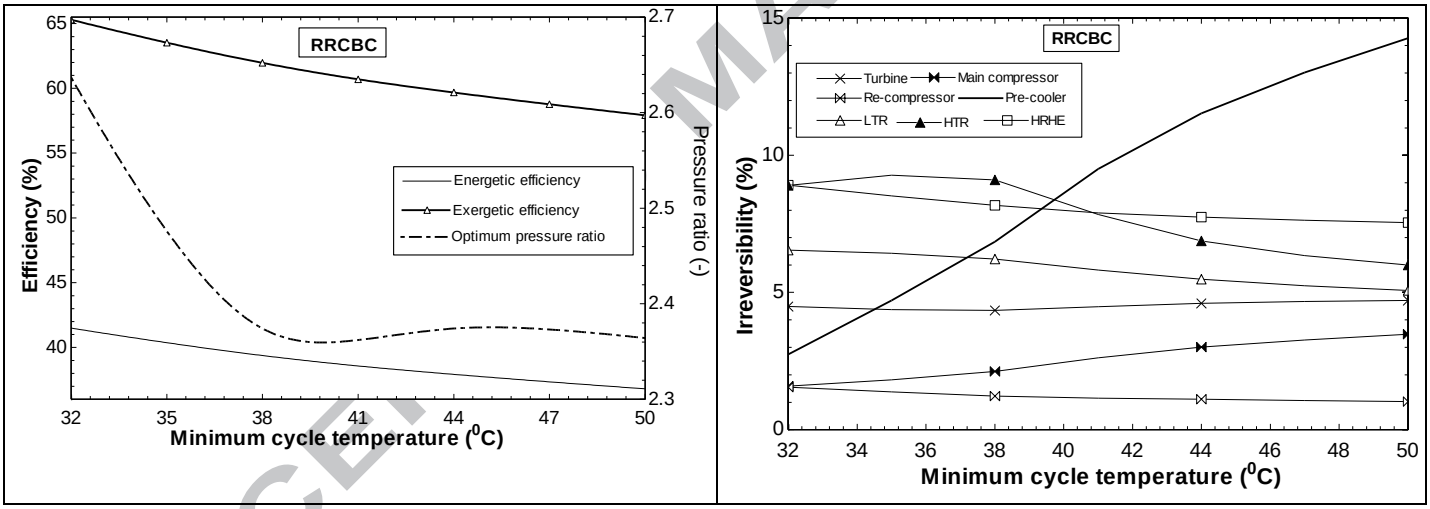


Fig. 4. Effect of minimum cycle (compressor inlet) temperature (T_5) on **(a)** cycle performance and optimum cycle pressure ratio ($r_{p,\,opt}$); **(b)** component irreversibility.

ACCEPTED MANUSCRIPT

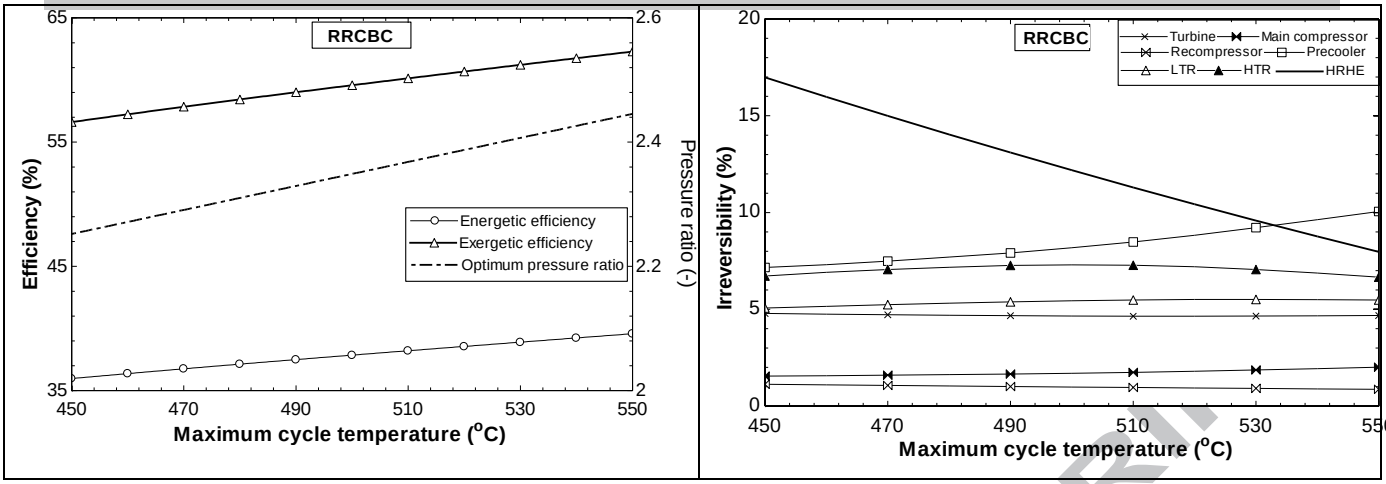


Fig. 5. Effect of maximum cycle (turbine inlet) temperature (T_1) on **(a)** cycle efficiency and optimum cycle pressure ratio ($r_{p,\,opt}$); **(b)** component irreversibility.

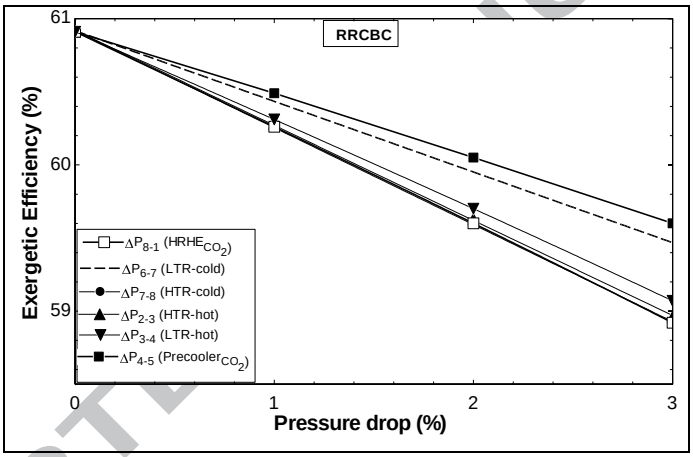


Fig. 6. Effects of component pressure drops on cycle exergetic efficiency.

ACCEPTED MANUSCRIPT

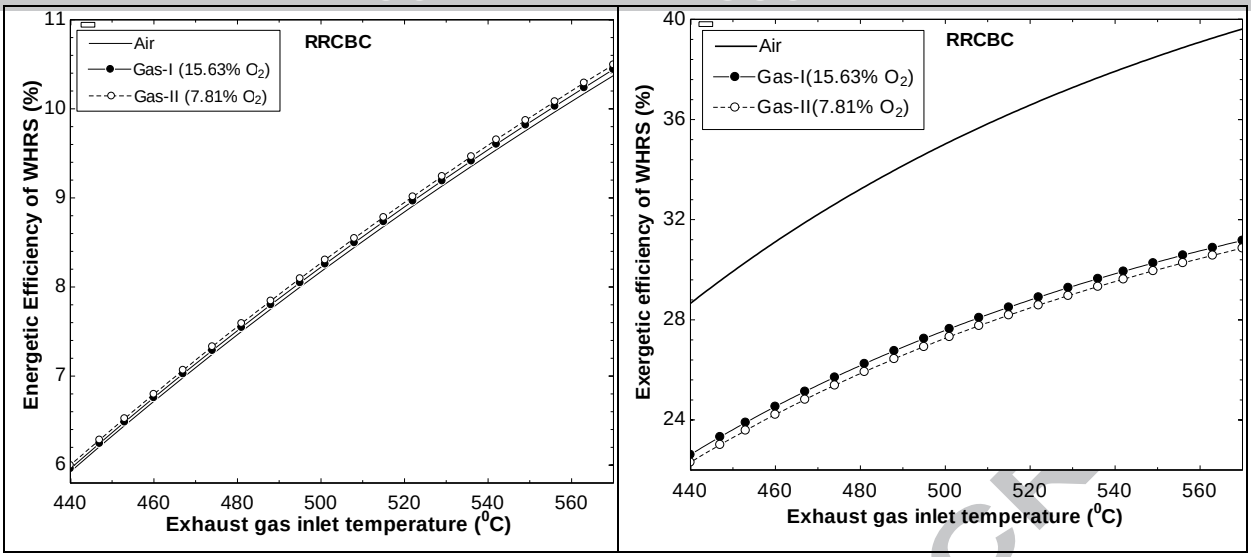


Fig. 7. Effect of gas compositions and gas inlet temperature on **(a)** the first law efficiency **(b)** the second law efficiency, of *WHRS* [Standard condition, T_1=723.15 K; gas compositions: air; gas-I: 15.63% O_2; gas-II: 7.81% O_2].

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1 Broad specifications of the shipboard marine gas turbine Ref. [49].

Term	ISO condition	Shipboard scenario
Fuel	*	LSHSD
Power Output (MW)	25.1	20.6
Power Level (%)	100%	100%
Thermal Efficiency (%)	38	35
Exhaust Mass Flow Rate (kg/sec)	70.5	63.6
Gas Temperature	839.15 K (566 °C)	845.15 K(572 °C)
Gas Pressure(MPa)	-	0.1084 (1.084 bar)
Intake Pressure Loss(Inch, WG)	Nil	8.5 (0.0212 bar)
Uptake Pressure Loss(Inch, WG)	Nil	24 (0.0597 bar)
Ambient Conditions	298.15 K (15 °C), 1 bar, 60% RH	308.15 K (35°C), 1 bar, 60% RH

*As per GE's fuel specifications for aeroderivative gas turbines.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 2 Details of exhaust gas composition scenarios.

	Mass fraction of combustion products, $y_{i\,(\%)}$								
Scenari	*Ar*	N_2	O_2	CO_2	SO_2	*CO*	*NO*	NO_2	*Othe*
Air	0.930	78.080	20.950	-	-	-	-	-	0.01
Gas-I	1.292	75.766	15.629	7.232	0.050	0.001	0.015	0.015	-
Gas-II	1.292	76.766	7.815	14.053	0.050	0.001	0.012	0.012	-

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3 Standard operating and design condition.

Operating and design data	Set value
Shipboard gas turbine power (%)	100
Gas composition	Ideal gas air or as specified
Gas mass flow rate (kg/s)	63.6
Gas inlet temperature (K)	845.15 (572 °C)
Gas inlet pressure (MPa)	0.1084
SC-CO_2 turbine inlet or maximum cycle temperature (K)	823.15 (550 °C)
SC-CO_2 compressor inlet or minimum cycle temperature (K)	313.15 (40 °C)
SC-CO_2 compressor discharge or maximum cycle	20
SC-CO_2 compressor inlet or minimum cycle pressure (MPa)	>7.37 MPa
PPTD in HRHE (hot side)	22
PPTD in pre-cooler (hot side)	15
Cooling water (sea water) inlet temperature (K)	303.15 (30 °C)
Component pressure loss in heat exchangers (%)	Nil
SC-CO_2 turbine isentropic efficiency (%)	90
SC-CO_2 compressor isentropic efficiency (%)	85
Effectiveness of HTR (%)*	86
Effectiveness of LTR (%)*	86
Effectiveness of HRHE (%)	90

(*) A conservative value assumed as per the published literature [14].

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 4 Comparison of results with the published literature.

Design parameter	Literature [9]		Literature [13]
	Con. design* (η_I = 45.27%)	B.E. design* (η_I = 47.36%)	Prototype design (η_I = 42.4%)
Maximum cycle temperature	550 ^{0}C	550 ^{0}C	480 ^{0}C
Minimum cycle temperature	32 ^{0}C	32 ^{0}C	35 ^{0}C
Maximum cycle pressure	20 MPa	20 MPa	20 MPa
Minimum cycle pressure	-	-	-
Cycle pressure ratio	2.6	2.6	2.35
MC isentropic efficiency	89%	95.5%	88%
RC isentropic efficiency	89%	94.8%	88%
Turbine isentropic efficiency	90%	92.9%	93%
HTR effectiveness	96.3%	96.3%	98%
LTR effectiveness	92.1%	92.1%	90%
Pressure drop	(heat exchangers = 130 kPa)	(heat exchangers = 130 kPa)	(Heat source=1%, HTR=0.4%, LTR=1.2%, Precooler=1%)
Present model results:			
Thermal Efficiency, η_I	46.56%#	48.45%#	42.41%
Deviation (%)	(+)2.8%	(+)2.3%	(+)0.02%

*Con. = Conservative design; *B.E. = Best-Estimate design (based on turbomachinery isentropic efficiencies assumed) [9]; #without pressure drops in heat exchangers.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 5 Thermodynamic properties at defined state points for supercritical CO_2 regenerative recompression Brayton cycle, at standard operating and design condition.

State point	Pressure (MPa)	Temperature (^{0}C)	Enthalpy (kJ/kg)	Entropy (kJ/kg-K)	Specific exergy flow (kJ/kg)
1	20	550.0	528.3	0.002	528.1
*2s**	8	429.2	394.6	0.002	394.4
2	8	440.7	408.0	0.021	401.9
3	8	251.3	190.9	-0.332	293.6
4	8	124.4	44.29	-0.653	245.8
5	8	40.0	-103.8	-1.081	229.7
*6s**	20	103.8	-71.86	-1.081	261.6
6	20	105.5	-67.91	-1.071	262.4
*7s**	20	213.9	110.1	-0.653	311.6
7	20	220.0	118.2	-0.636	314.6
8	20	393.1	335.3	-0.258	415.2

(*): isentropic or ideal state

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 6 Exergy balance for RRCBC at standard operating and design condition, without pressure drop, and optimum cycle pressure ratio=2.47.

Exergy	kJ/kg	% (of exergy input)	Irreversibility	kJ/kg	% (of exergy input)	% (of total irreversibility)
Exergy input to the cycle	122.7	100.0	HRHE	9.8	8.0	21.1
			Precooler	12.3	10.1	26.7
			HTR	8.2	6.7	17.7
			LTR	6.7	5.5	14.5
Exergy recovered through Turbine	119.0		Turbine	5.7	4.7	12.4
Exergy input to Main-compressor	27.3		Main compressor	2.5	2.0	5.3
Exergy input to Re-compressor	15.3		Recompressor	1.1	0.9	2.3
Net exergy output	76.4	62.3	Total Irreversibility	46.3	37.7	100.0

ACCEPTED

ACCEPTED MANUSCRIPT

Table 7 Optimized performance results at standard operating and design condition at different pressure drop fractions.

Pressure drop	$r_{p,opt}$ (*)	$\eta_{I\,(RRCBC)}$	$\eta_{II\,(RRCBC)}$	$\eta_{I\,(CC)}$	$\eta_{I\,(WHRS)}$	$\eta_{II\,(WHRS)}$	Net power	BWR	x
(%)	(-)	(%)	(%)	(%)	(%)	(%)	(kW)	(-)	(-)
0	2.47	39.6	62.3	45.1	8.6	32.9	5148	0.36	0.212
1	2.49	37.4	58.8	44.4	7.9	30.2	4721	0.38	0.206
3	2.63	32.2	50.7	42.9	6.5	24.6	3842	0.46	0.196

(*): $r_{p,opt}$ corresponds to maximum cycle efficiency.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

HIGHLIGHTS

- A shipboard gas turbine exhaust driven CO_2 recompression power cycle is proposed.
- Effect of key design parameters on energetic & exergetic performance presented.
- HRHE and HTR are the most critical component with respect to pressure-drop.
- Heat exchangers involve about 4 times irreversibility than the turbo-machinery.
- Gas turbine exhaust composition, temperature and load influence WHRS performance.

ACCEPTED MANUSCRIPT

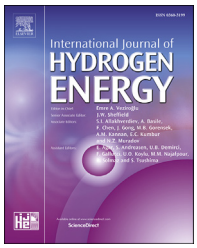
INTERNATIONAL JOURNAL OF HYDROGEN ENERGY XXX (2017) 1–11



Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he



Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid

Nikita Gupta[*], *Rachana Garg*

Department of Electrical Engineering, Delhi Technological University, Delhi, India

ARTICLE INFO

Article history:
Received 6 March 2017
Received in revised form
6 May 2017
Accepted 14 May 2017
Available online xxx

Keywords:
Solar photovoltaic system
Control algorithms
PV inverter
Fuzzy logic control
Total harmonic distortion

ABSTRACT

Solar Photovoltaic (SPV) cells and hydrogen fuel cells are green and efficient Distributed Energy Resources (DERs) with minimal environmental impact. Integration of DERs with conventional grid through inverters has proven to be an accepted technique for a continuous and good quality power to end users. For an efficient use of an inverter, its control algorithm plays an important role. In this paper, asymmetrical Fuzzy Logic Control (FLC) algorithm is designed, developed and simulated for a SPV system connected to the grid. In proposed asymmetrical FLC algorithm, the fuzzy functions of error input close to zero are used for fine-tuning, and the error input away from zero are used for coarse tuning. The effectiveness of the proposed control algorithm is demonstrated for controlling the PV inverter in unity power factor mode while maintaining the power quality standards and load balancing. Further, the system steady-state and transient response with the asymmetrical FLC algorithm is analyzed for linear as well as non-linear loads using MATLAB along with Simulink toolbox. Finally, a comparative analysis of asymmetrical FLC algorithm with other algorithms is presented to establish the superiority of the proposed algorithm.

© 2017 Hydrogen Energy Publications LLC. Published by Elsevier Ltd. All rights reserved.

Introduction

Energy is an essential ingredient for the process operation and its sustainability. Electrical energy is a form of energy i.e. easy to transport, utilize, process and control, is preferred in all walks of life such as industries, communication, transportation, domestic appliances, etc. With the depletion of fossil fuels and the environment pollution, renewable energy sources such as solar, hydrogen fuel cell, wind, hydro, etc. have taken center stage for electrical power generation [1,2]. Solar Photovoltaic cells and hydrogen fuel cells are green energy efficient sources with minimal environmental impact [3]. Due to the abundance of sunlight, solar thermal and SPV have gained widespread acceptance [4–13]. SPV technology directly transforms solar light energy to electrical energy. The electrical power extracted from SPV is not constant, and hence DC–DC converter is required to stabilize the DC output of the SPV and to extract maximum power output [14–16]. DC–AC inverter is added for supplying power to the AC loads and to meet the supply demand equilibrium. Inverter control in a SPV system connected to grid plays a vital role in processing, control and synchronization of SPV generated power with the grid. PV inverter control is based on the estimation of active, reactive powers and harmonic currents [17–21]. Further, the required PV inverter performance depends on control algorithm used for generation of reference currents which in turn generate the gate pulses for PV inverters' switches.

* *Corresponding author.*
E-mail address: guptanikita08@gmail.com (N. Gupta).

http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

0360-3199/© 2017 Hydrogen Energy Publications LLC. Published by Elsevier Ltd. All rights reserved.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

Nomenclature

DER	Distributed Energy Resource
FLC	Fuzzy Logic Control
SPV	Solar Photovoltaic
THD	Total Harmonic Distortion
SRF	Synchronous Reference Frame
IRP	Instantaneous Reactive Power
$I_{PV(cell)}$	SPV cell output current
κ	Boltzmann constant
q	Elementary charge (1.602×10^{-19}C)
A	Diode Ideality Factor
R_s	series resistance (Ω) of solar cell
R_p	shunt resistance (Ω) of solar cell
E	Operating irradiation
E_{ref}	SPV cell reference irradiation
T_{ref}	SPV cell reference temperature, 278 K
T	SPV cell operating temperature, in K
k_i	I_{SC} temperature coefficient
I_{ph}	Photocurrent
V_{oc}	Open circuit voltage
I_{sc}	Short circuit current
L	Inductor of boost converter
C	Capacitor of boost converter
D	Duty cycle of boost converter
V_i	Input voltage of boost converter
V_o	Output voltage of boost converter
Δi_i	Input current ripple of boost converter
Δv	Output voltage ripple of boost converter
I_o	Output current of boost converter
f_s	Switching frequency
L_f	Filter Inductance
R_f	Internal resistance of the filter inductor
i_s	Grid Currents
v_s	Grid Voltages
v_c	Inverter Voltages
P	Active Power
Q	Reactive Power
IGBT	Insulated-Gate Bipolar Transistor
PWM	Pulse Width Modulation
HCC	Hysteresis Current Control
error	Crisp input to Fuzzification ($V^*_{dc} - V_{dc}$)
derror	Crisp input to Fuzzification ($\frac{d}{dt}error$)
V^*_{dc}	Reference DC link voltage
V_{dc}	DC link voltage
i_{loss}	Output of FLC
defuzzifier	Defuzzification Operator
x	Fuzzy i_{loss} Output
UPF	Unity Power Factor
i_c	Inverter Currents
v_l	Load Voltages
i_l	Load Currents
PCC	Point of Common Coupling

The photon energy input to SPV module and environmental temperature are intermittent in nature [2,4]. Also, the type and quantum of loads connected to the source depend upon the customers' requirements. Under all these conditions, utilization of SPV module efficiently needs dynamic output voltage and current to be tracked and optimized using appropriate control algorithm. In literature, many control algorithms have been developed for PV inverters. Broadly they can be classified as conventional, intelligent and adaptive algorithms. Some important conventional control algorithms are Synchronous Reference Frame (SRF) theory, unit template, Instantaneous Reactive Power (IRP) theory, conductance fryze, etc. [22–27]. Conventional control algorithms are not best suitable for nonlinear variations as the overshoots and long settling time in their response are inevitable. To overcome these problems, a PV inverter having an intelligent or adaptive algorithm can result in better accuracy with wide adjustment range of its internal parameters along with fast dynamic response [28–43].

The FLC algorithm is one such intelligent control algorithm that can take into account the uncertainty and vagueness factors involved in the input and load requirements [28,29]. Soufi et al. [30] developed an improved maximum power point tracking method based Fuzzy-Particle swarm optimization algorithm for solar panel with faster tracking speed and zero oscillations. Garcia et al. [31] presented and evaluated a novel hourly energy management system for a stand-alone hybrid system and showed that the proposed system based on fuzzy logic ensures the optimization of the utilization cost. Bouafia et al. [32] designed a PI and FLC controller for three-phase PWM rectifier and showed that the fuzzy is better in steady and transient states. Cao and Wang [33] designed a single-phase grid-connected SPV system based on Fuzzy-PID controller to control the switch of the DC–DC converter. They obtained SPV system with improved the power factor and lesser harmonics.

Yatak and Bay [34] showed the application of FLC for the three-phase inverter of SPV system tied to the grid and presented that fuzzy based controller leads to the precise tracing of reference currents. Hannan et al. [35] showed the real time testing of FLC for grid-connected three-phase PV inverter. They presented the results for capability of fuzzy to generate quality PV power while maintaining unity power factor. Sakhare et al. [36] applied the fuzzy logic control for fuel cell application. They used the FLC to control DC–DC converter and DC–AC converter and got the inference that implementation of fuzzy based controller is cost effective. Vigneysh and Kumarappan [37] presented the simulation comparison of FLC controller and PI controller in order to maintain the microgrid system voltage and frequency. Thao and Uchida [38–40] showed the detailed application of FLC in converter as well as inverter control in a grid connected SPV system. Authors demonstrated a simple single-level control method suitable for a small SPV plant and also coordinated control strategy for a large-scale SPV farm. They have focused on frequency and power flow control of grid connected SPV system.

The present literature regarding grid connected SPV system focusses mainly on AC voltage generation by PV inverter and synchronization with grid using FLC with parametric variation such as atmospheric uncertainties and AC load change. However, the multi-objective advantages of PV inverter are still unexplored that can be utilized and improved along with frequency and power flow control in steady state and transient conditions.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

The uncertainties in SPV system connected to grid are highly complex and random in nature, symmetrical fuzzy functions of FLC do not give satisfactory results to meet the requirements. For such conditions, asymmetrical fuzzy functions can be better suited. In this paper, the asymmetrical fuzzy logic control algorithm is proposed for SPV system connected to the grid. The steady-state and transient performance of the algorithm is studied along with a detailed analysis of its multifunctional capabilities like active filtering, reactive power compensation and load balancing for a different type of load conditions. A comparative analysis of various algorithms is then carried out to establish the superiority of the proposed algorithm. The system under consideration is simulated using MATLAB along with Simulink toolbox.

Problem formulation

Non-linear variations in input parameter (photon energy, temperature) and demand (loads) in a SPV system connected to grid needs a control algorithm which can take into account the uncertainties in the parameters of interest. FLC algorithm, a soft computing tool meets this situation. However, the performance of FLC algorithm depends heavily on the defining of its membership functions and its fuzzy rules. Further, both photon energy and the load are highly complex and random in nature. Asymmetrical input membership functions of FLC and their carefully designed fuzzy rules can well control and improve the dynamic response of a system.

In the present study, the authors have considered a 10 kW SPV system connected to the grid as shown in Fig. 1. SPV system connected to grid consists of SPV array, followed by DC–DC boost converter to raise the voltage to the desired voltage level and perform with maximum power tracking function using perturb and observe algorithm [15]. DC-link is connected to PV inverter which in the present work is a voltage source inverter. The inverter is controlled using the asymmetrical FLC algorithm, and its steady-state and transient behavior is analyzed. Extensive studies of the PV inverter ancillary services using the asymmetrical FLC algorithm are demonstrated using detailed simulation results. Also, a comparison is presented between the performances of the proposed algorithm with other existing algorithms.

Mathematical model of SPV system

The block diagram representation shown in the lower part of Fig. 1 comprises three main sections of a SPV system that are SPV array, DC–DC boost converter and PV inverter. The design and mathematical model of each section are briefed in the following subsection studied and presented by the authors in Ref. [4].

SPV array

A SPV cell, an active transducer, converts the photon energy of sunlight directly into electricity. When exposed to sunlight (or another intense light source), a solar cell produces almost constant open circuit voltage depending upon the cell material while the current ($I_{PV(cell)}$) flow to connected load is proportional to light energy (photons) and is given by Refs. [4,5].

$$I_{PV(cell)} = I_{ph} - I_o\left[\exp\left\{\frac{q(V_{PV} + I_{PV}R_s)}{AKT}\right\} - 1\right] - \frac{V_{PV} + I_{PV}R_s}{R_p} \quad (1)$$

where I_{ph} is the photocurrent, κ is constant, called the Boltzmann constant, q is the elementary charge (1.602×10^{-19}C), A is the diode ideality factor, R_s is the series resistance (Ω), and R_p is the shunt resistance (Ω) of the solar cell. The equation for

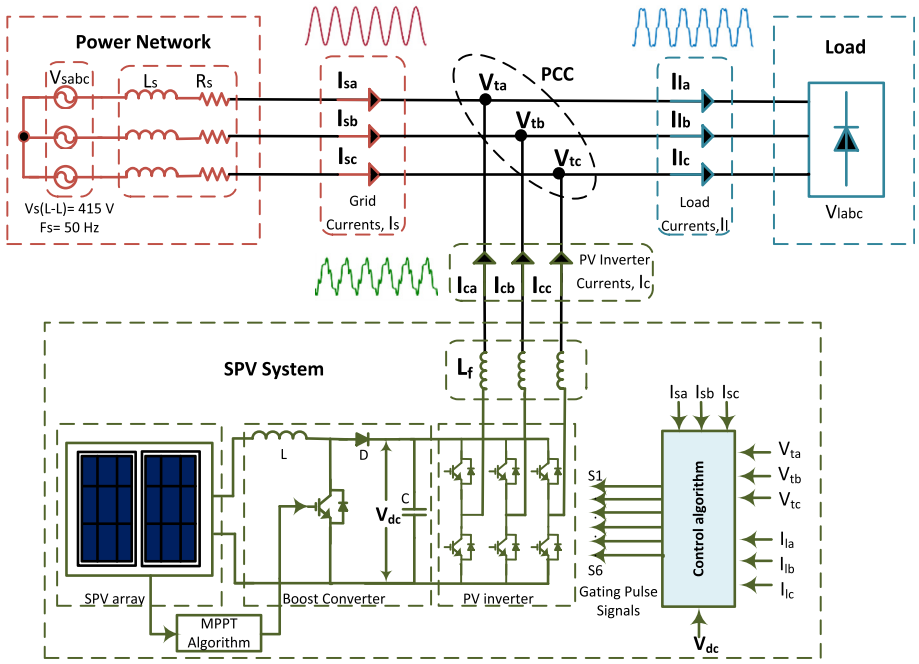


Fig. 1 – Block diagram of the SPV system connected to the grid.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

the SPV array is obtained by modifying the above cell equation by inserting the indices n and m where n and m are number of cells in series and parallel, respectively, and is given by

$$I_{PV(array)} = mI_{ph} - mI_o\left[\exp\left\{\frac{q\left(V_{PV(array)} + I_{PV(array)}R_s\left(\frac{n}{m}\right)\right)}{nAKT}\right\} - 1\right] - \frac{V_{PV(array)} + I_{PV(array)}R_s\left(\frac{n}{m}\right)}{R_p\left(\frac{n}{m}\right)} \quad (2)$$

The photocurrent I_{ph}, depends upon parameters, such as temperature and solar irradiance, and is defined by

$$I_{ph} = I_{sc}\frac{E}{E_{ref}}\left[1 + (T - T_{ref})*k_i\right] \quad (3)$$

where E is operating irradiation, E_{ref} is SPV cell reference irradiation, T is the SPV cell operating temperature, K, T_{ref} is SPV cell reference temperature of 278 K, and k_i is constant related to temperature co-efficient due to I_{SC}. Fig. 2 depicts the effect of irradiation level and temperature on current vs voltage and power vs voltage curve of the given SPV array. The array used in simulation work has a peak capacity of 10 kW, and is designed for V_{oc} of 403 V and I_{sc} of 34.5 A [4].

DC–DC boost converter

In the designed system, DC–DC boost converter (Fig. 1) is used to obtain and regulate the required dc bus voltage. The passive elements, an inductor (L) and capacitor (C), of the converter, are computed using Equations (4) and (5) respectively [4,14]:

$$L = V_i(V_o - V_i)/\Delta i_i f_s V_o \quad (4)$$

$$C = I_o D/\Delta v f_s \quad (5)$$

$$D = 1 - (V_i/V_o) \quad (6)$$

where V_i and V_o are the input and output voltages, respectively of the converter, Δi_i is input current ripple taken as 10% of input current, f_s is the switching frequency of IGBTs switches, Δv is output voltage ripple taken as 3% of the boost converter output voltage, I_o is the output current and D is the duty cycle of boost converter [14].

PV inverter

A three phase PV inverter is used to convert DC voltage to AC voltage. The two-level PV inverter consists of six IGBTs as shown in Fig. 1. In this paper, a 10 kW three phase 415 V inverter [4,14] is designed with 0.9 power factor load.

PV inverter algorithm

The primary objective of the control algorithm of PV inverter is to estimate the reference current signals using feedback with different controlling parameters (i_l, v_t, i_s, and v_{dc}) and to generate the gate triggering pulses for the inverters' IGBTs [22,23]. A generalized block diagram of the closed loop control of PV inverter is shown in Fig. 3 comprising of three main blocks namely voltage control, network synchronization control and gating pulse generation control. The inverter can be made to operate in two control modes: unity power factor and zero voltage regulation. The choice of control strategy of PV inverter is determined by the type of control intended, systems' stability and power quality requirement.

The voltage control block is responsible for voltage control across DC-link capacitor shown in the lower part of Fig. 1, which depends on the energy balance of the power received by the PV inverter and the power delivered by it [19,23]. Voltage control loop uses controllers for forcing the actual quantities to follow the reference quantities. The controller processes the error signal ($V^*_{dc} - V_{dc}$) to generate a new amplitude of current component i_{loss} that adds to the reference current signals for any energy imbalance in DC bus bar. Conventional voltage control loop uses Proportional–Integral

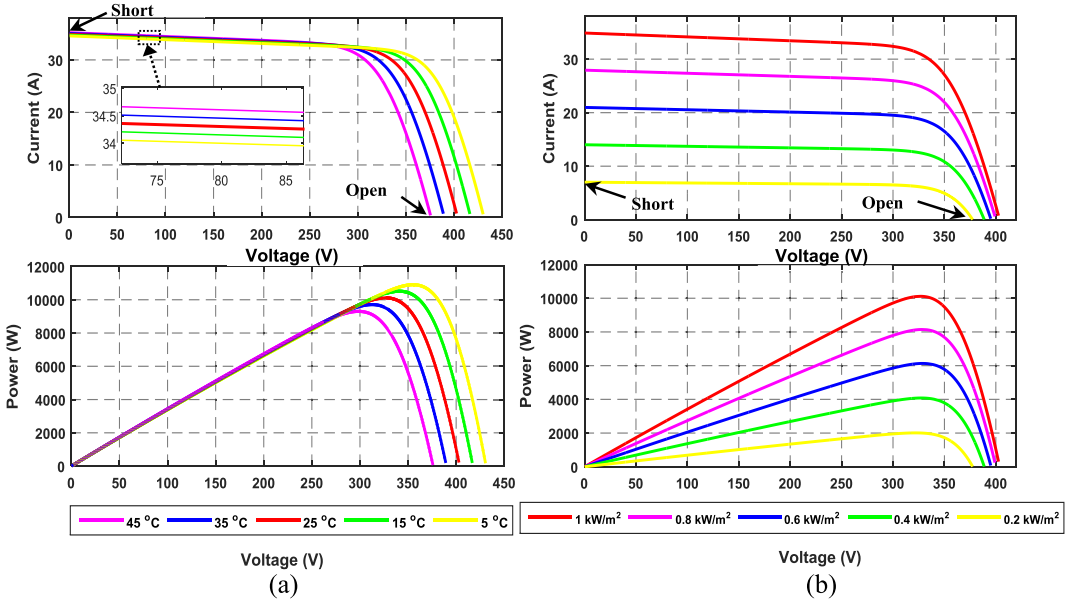


Fig. 2 – Current Vs Voltage and Power Vs Voltage curves of SPV array used in proposed system model. Under a) Different temperature levels, and b) Different irradiation level.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

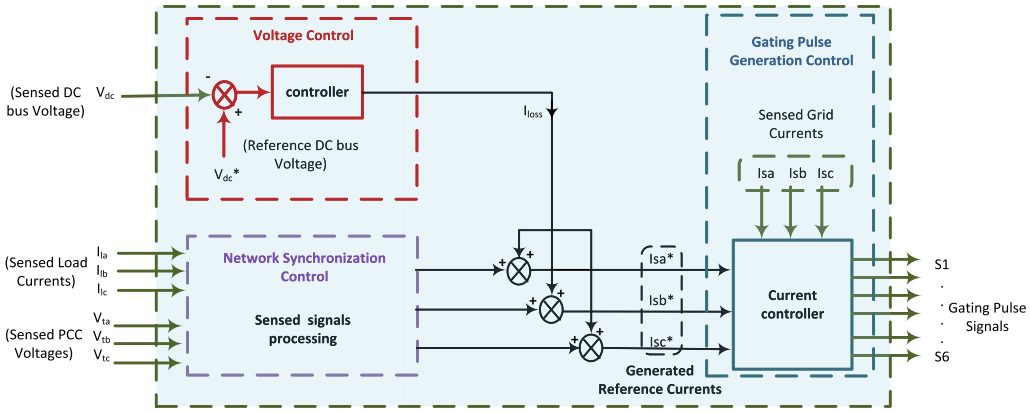


Fig. 3 – Block diagram of control algorithm of PV inverter.

(PI) controller, which is a simple first order controller. In the present study, the asymmetric FLC is introduced in place of PI to control the voltage across the DC-link capacitor.

Network synchronization control block is a basic current controller block in which signal processing is carried out, and fundamental component of the control signal is extracted [22,23]. The inverter output voltage is the vector sum of the AC grid voltage and the series drop of the filter as shown in Fig. 4. Therefore, the network synchronization control block estimates the converter output voltage for a given power reference accounting the filter drop and maintains the power flow to follow the reference.

Various reference frames can be adapted for the current control viz. natural reference frame, stationary reference frame, and synchronous reference frame. The synchronous reference frame is used in the present work. The voltage equation of the equivalent PV inverter model in the stationary and rotating frame can be written as shown in Equations (7) and (8), respectively [20,21].

$$v_c = R_f\, i_s + L_f \frac{d(i_s)}{dt} + v_s \tag{7}$$

$$v_{cdq} = R_f\, i_{sdq} + L_f \frac{d(i_{sdq})}{dt} + v_{sdq} \tag{8}$$

where L_f is the filter inductance and R_f is the internal resistance of the inductor. i_s, v_s and v_c are grid currents, grid voltages and inverter voltages, respectively in stationary frame and i_{sdq}, v_{sdq} and v_{cdq} are grid currents, grid voltages and inverter voltages, respectively in synchronously rotating frame. Taking Laplace transform of Equation (8), the transfer function $T(s)$ is derived and presented in Equation (9)

$$T(s) = \frac{i_{sdq}(s)}{V_{cdq}(s) - V_{sdq}(s)} = \frac{1}{(s + jw)L_f + R_f} \tag{9}$$

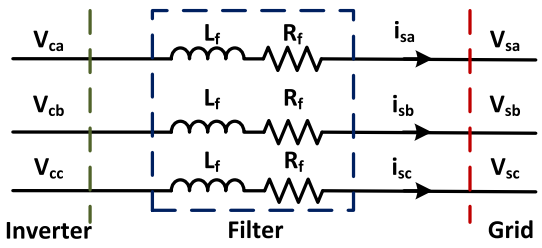


Fig. 4 – Space vector equivalent of PV inverter.

The grid voltage (v_{sdq}), is subtracted from the inverter voltage (v_{cdq}), for the required output current. PV inverter supplies the desired voltage, v_c with low harmonic distortion. Fundamental components obtained from network synchronization control block along with voltage control signal are used to calculate the reference current signals [22,23]. Reference current signals, thus obtained along with corresponding sensed current signals are used to generate the gating pulse signals for switching devices (IGBTs) of PV inverter in the gating pulse generation control. Pulse Width Modulation (PWM) or Hysteresis Current Control (HCC) technique can be used for gating pulse generation [23]. In the present work, HCC is used.

Proposed asymmetrical FLC algorithm

Conventional control algorithms are not suitable for nonlinear variations since their response comprises of overshoot and larger settling time. Fuzzy logic, an intelligent algorithm based on symmetrical membership functions gives a satisfactory solution for nonlinear systems with uncertainties. Unlike Boolean logic, it deals with problems that have fuzziness or vagueness [28]. The methodology of reasoning in fuzzy logic uses "IF… THEN…"statements or rules [29]. In general, membership functions are assigned to the input and output variables. Fuzzy logic membership functions can be of two types-symmetrical and asymmetrical.

In the present studies, asymmetrical membership functions are proposed since the parameters associated with SPV systems such as photon energy and the load are uncertain and random in nature. Asymmetrical fuzzy functions possess the fine tuning and coarse tuning features, and can provide better controlling and stability. Asymmetrical input membership function of FLC and its carefully designed fuzzy rules can well control and improve the dynamic response of a system. Fig. 5 shows the block diagram of asymmetrical FLC controller used in the voltage control of PV inverter in SPV system connected to the grid.

Fuzzification

The proposed asymmetrical FLC is used to regulate the DC link voltage across the capacitor, C. The sensed DC voltage, V_{dc} is

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

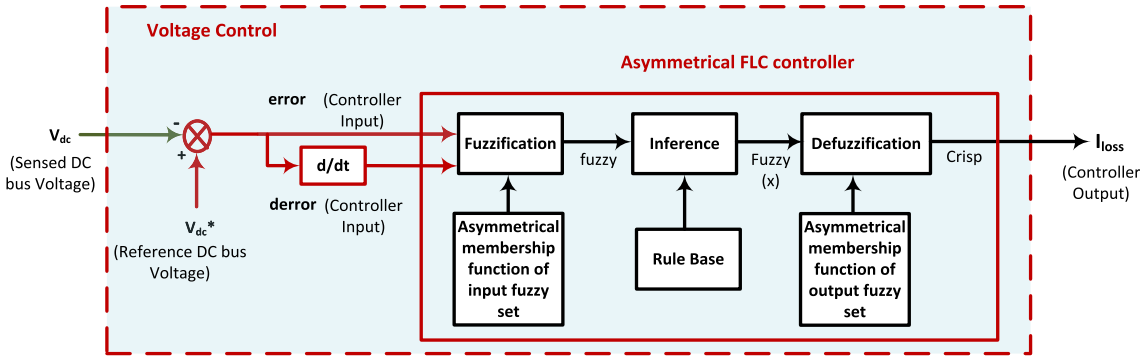


Fig. 5 – Block diagram of asymmetrical FLC controller applied for the voltage control of PV inverter.

compared with a reference value, V_{dc}^{*}(= 800 V) which generates the i_{loss} factor used for regulating active power component of current [29,42]. The controller consists of two input variables: *error* and *derror* (change in error), and one output variable i_{loss}.

$$error(N) = V_{dc}^{*} - V_{dc}(N) \tag{10}$$

$$derror(N) = error(N) - error(N-1) \tag{11}$$

where $error(N)$ is the error and $derror(N)$ is the change in error in the *N*th iteration. V_{dc}^{*} is the reference dc link voltage and $V_{dc}(N)$ is the sensed dc linked voltage in *N*th iteration [42].

Implementation of asymmetrical FLC is explained in the following subsections:

Fuzzy sets of input and output variables consist of seven sets of triangular membership functions as shown in Fig. 6. The membership functions are *–Hi*: negative high, *–Me*: negative medium, *–Lo*: negative low, *Ze*: zero, *+ Lo*: positive low, *+ Me*: positive medium and *+ Hi*: positive high [41]. In the asymmetrical fuzzy functions, the narrow span functions (*–Lo*, Z, *+ Lo*) are considered near zero for *error* and i_{loss} factor. These functions act like fine tuning. Wide span functions are considered for *–Hi*, *-Me*, *+ Me*, and *+ Hi*. These functions act like coarse tuning.

Each membership function has a membership value belonging to {0 1}. Study of the PI controller operation of the system is used to set the span of the input and output membership functions to train the FLC system. Initially, V_{dc} the voltage across the DC link capacitor may be zero, whereas, V_{dc}^{*} the reference voltage is 800 V, so, the maximum error can be ±800 V. Thus, the span of *error* functions considered is {–800 800} V for the DC-link voltage error. Also, for the DC-link voltage, the simulation results shows that the error value varies between the ranges of {–150 150} during transient situation. Hence, near zero error i.e. in the range {–150 150}, the frequency of error considered is high. Similarly, the span for the *derror* is set as {–1 1} and for the output i_{loss} is set as {–60 60} A.

Inference method

The main function of the inference mechanism is to process the fuzzy values of input according to the rules drawn by experts and based on the active membership functions in *error* and *derror* inputs. Asymmetrical FLC algorithm implemented consists of total 49 rules. An example of a rule is shown as [40,42,43]:

IF
"*error*" is +Lo (Positive Low)
"*derror*" is –Lo (Negative Low) (12)
THEN
"i_{loss}" is Ze (Zero)

Fig. 7 shows the rule base of asymmetrical FLC algorithm both in a table and in a surface plot. With the help of tabular and surface view, the control rules of the fuzzy logic system are changed so that control action varies with respect to the rules. The rule base matrix provides rules for all possible combinations of the membership functions for *error* and *derror* in line with Equation (12). Thus, using information from the rule base, the rule, and its certainty is determined by the inference mechanism. Mamdani's MAX–MIN method for inference of FLC algorithm. The output membership function of each rule is given by minimum operator whereas collective fuzzy output (*x*) is provided by maximum operator [29].

Defuzzification

The output of inference method is a fuzzy set. Defuzzification is used to obtain required crisp control output $z_o(= i_{loss})$ and is defined by

$$z_o = defuzzifier\ (x) \tag{13}$$

where *x* is aggregate output and *defuzzifier* is defined as defuzzification operator [29]. In the present work center of the area is used as defuzzification operator and is given by the following equation:

$$x = \frac{\sum_{i=1}^{n}\mu(x_i)x_i}{\sum_{i=1}^{n}\mu x_i} \tag{14}$$

where μx_i is the activation degree on rule *i*, x_i is the center of the Max–Min composition of the output membership functions and *x* is the required output i_{loss}.

Simulation results

Steady-state and dynamic performance of asymmetrical FLC

PV inverter in unity power factor (UPF) mode should act as a VAR generator and supply the reactive power demand of load

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

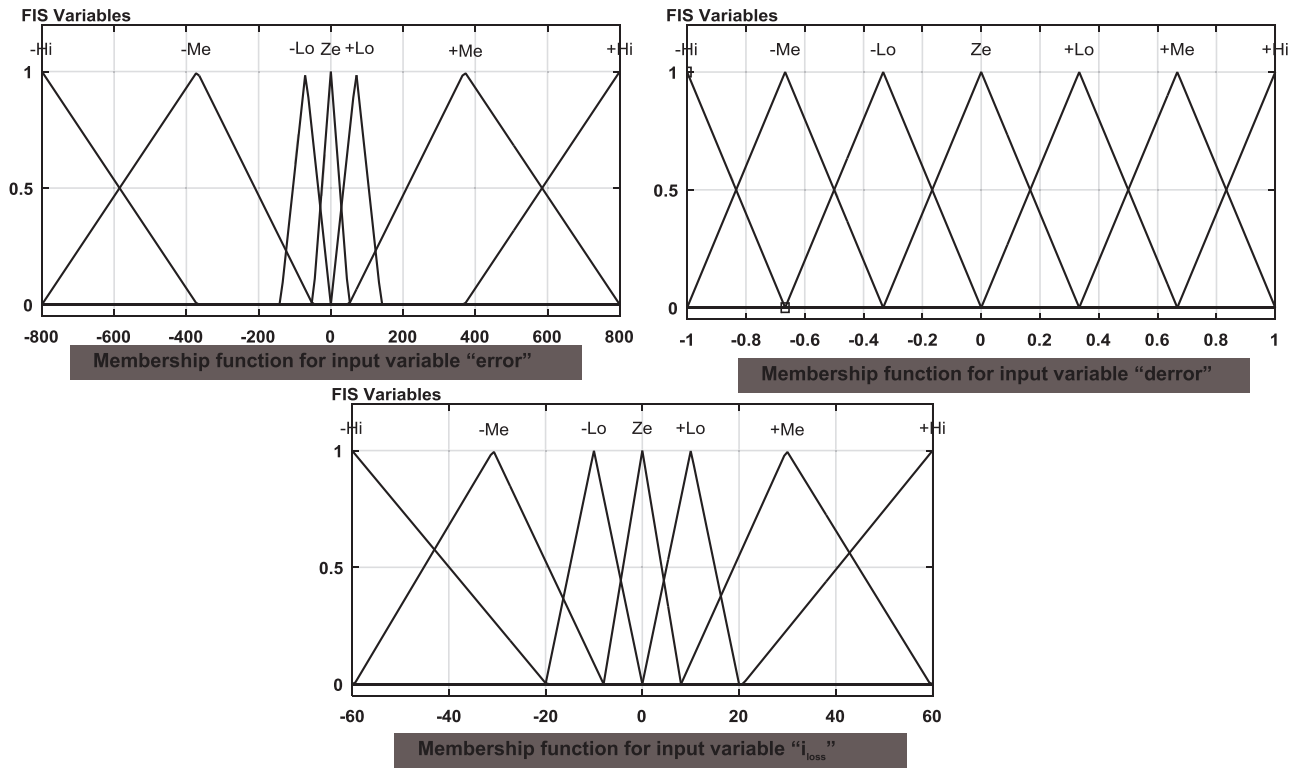


Fig. 6 – Membership functions for input variable '*error*', '*derror*' and output variable 'i_{loss}'.

locally. Fig. 8 shows the steady-state performance of the proposed algorithm. Fig. 8 (a)–(c) show the voltage and current waveforms of the grid, inverter, and load respectively. For testing the UPF control of asymmetrical FLC, a highly inductive linear load with power factor 0.6 was taken. It can be seen from Fig. 8 (a), grid current remains in phase with grid voltage. Also, total THD in grid current is 3.04%, which is well within IEEE 519-2014 standard [44].

Fig. 8 (d)–(f) show the active and reactive power supplied by the grid and the inverter to the load respectively. During steady-state, active and reactive power demand of load is 15 kW and 20 kVAR (lagging) respectively. The SPV source provides the active power of 10 kW, and the grid supplies the rest 5 kW. From Fig. 8 (d)–(f) it can be seen that complete reactive power demand of load is supplied by PV inverter locally, thereby, reducing reactive power supplied by the grid to zero. Thus, the asymmetrical FLC algorithm is efficient in maintaining unity power factor mode of operation along with power flow balance in a SPV system connected to the grid.

The PV inverter should act as a load balancer for single phasing in a three phase system. Fig. 9 (a) and (b) shows the response of asymmetrical FLC controlled PV inverter for load balancing action. To realize single phasing of load, phase '*a*' of a three phase highly inductive non-linear load is short-circuited. Asymmetrical FLC takes appropriate action and grid current in all three phases is settled down to the new value of 8.4 A (RMS)/phase from its old value 26.5 A (RMS)/phase. The PV inverter current for the phase '*a*' is found to be

error \ derror	-Hi	-Me	-Lo	Ze	+Lo	+Me	+Hi
-Hi	-Hi	-Hi	-Hi	-Me	-Me	-Lo	Ze
-Me	-Hi	-Hi	-Me	-Me	-Lo	Ze	+Lo
-Lo	-Hi	-Me	-Me	-Lo	Ze	+Lo	+Me
Ze	-Me	-Me	-Lo	Ze	+Lo	+Me	+Me
+Lo	-Me	-Lo	Ze	+Lo	+Me	+Me	+Hi
+Me	-Lo	Ze	+Lo	+Me	+Me	+Hi	+Hi
+Hi	Ze	+Lo	+Me	+Me	+Hi	+Hi	+Hi

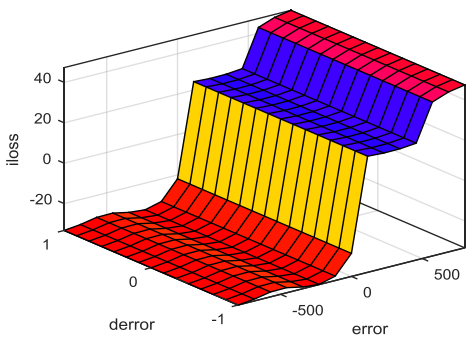


Fig. 7 – Rule base for computing 'i_{loss}' component in tabular and surface plot forms.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

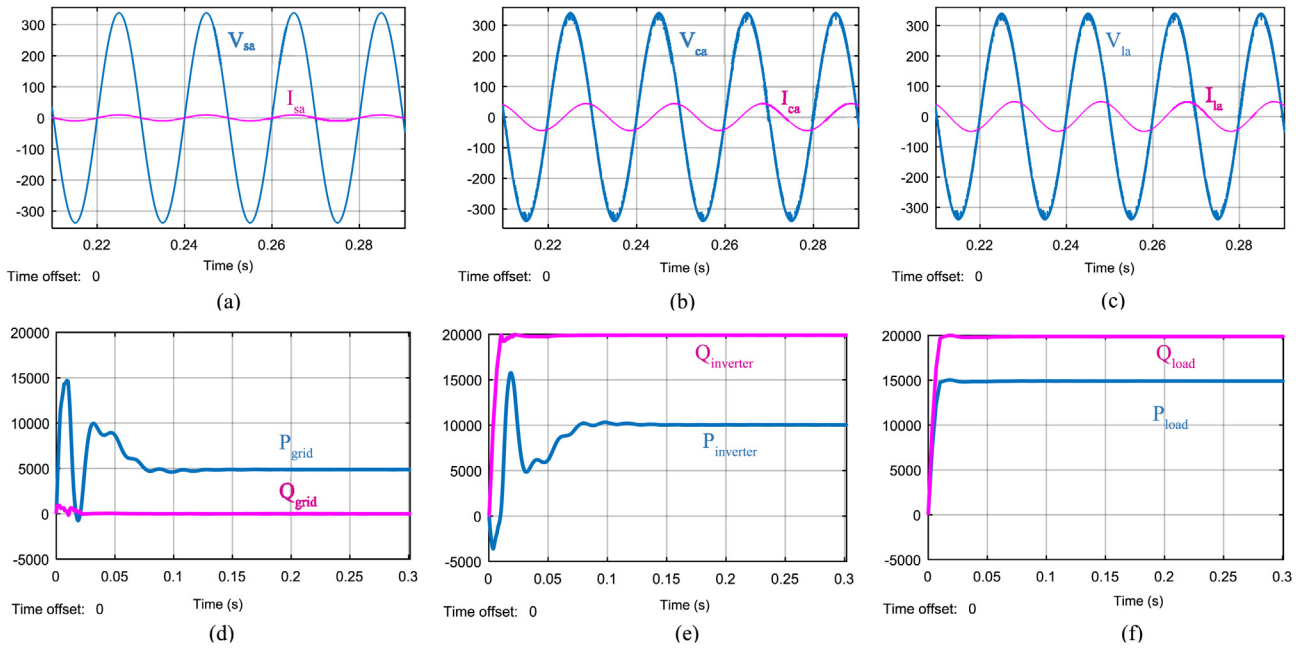


Fig. 8 – Steady-state performance of asymmetrical FLC algorithm. v_{sa} = 338 V (RMS), i_{sa} = 10 A (RMS), v_{ca} = 338 V (RMS), i_{ca} = 44 A (RMS), v_{la} = 338 V (RMS) and i_{la} = 49 A (RMS). P_{grid} = 5 kW, Q_{grid} = 0 kVA, $P_{inverter}$ = 10 kW, $Q_{inverter}$ = 20 kVA, P_{load} = 15 kW and Q_{load} = 20 kVA.

equal and opposite to the grid current, as the load is zero in this phase. The PV inverter draws current i_{sa} from the grid. Hence, grid sees a combination of load and PV inverter as a three-phase balanced load. Also, i_{sa} is in phase with v_{sa}, which is the requirement of UPF operation. Asymmetrical FLC is able to perform active filtering as well. THD of grid current is maintained at 1.07%, even when THD of connected load is 16.34%, which is well within the IEEE 519-2014 standard. Thus, the Asymmetrical FLC algorithm effectively carries out load balancing and harmonic compensation.

Comparison of asymmetrical FLC algorithm with other algorithms

Fig. 10 (a) and (b) show the THD % of the PCC (Point of common coupling) voltage and grid current respectively, for SRF algorithm, conductance fryze algorithm, unit template algorithm, adaptive algorithm and proposed asymmetrical FLC algorithm. THD of the PCC voltage and grid current were observed under three conditions, i.e. steady-state, during load increase and during load unbalancing. Asymmetrical FLC algorithm

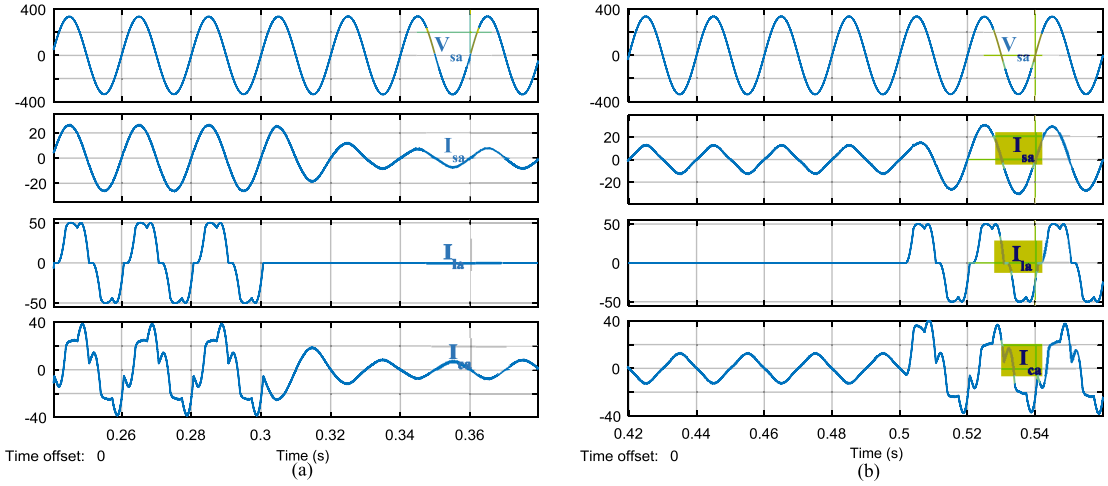


Fig. 9 – Dynamic performance of asymmetrical FLC algorithm to phase *a* load disconnection, v_{sa} = 338 V (RMS), i_{sa} = 26.50 A (RMS) to 8.4 A (RMS), i_{la} = 50 A (RMS) to zero and i_{ca} = 40 A (RMS) to 8 A (RMS). v_{sa} = 338 V (RMS), i_{sa} = 8.4 A (RMS) to 26.50 A (RMS), i_{la} = zero to 50 A (RMS) and i_{ca} = 8 A (RMS) to 40 A (RMS).

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

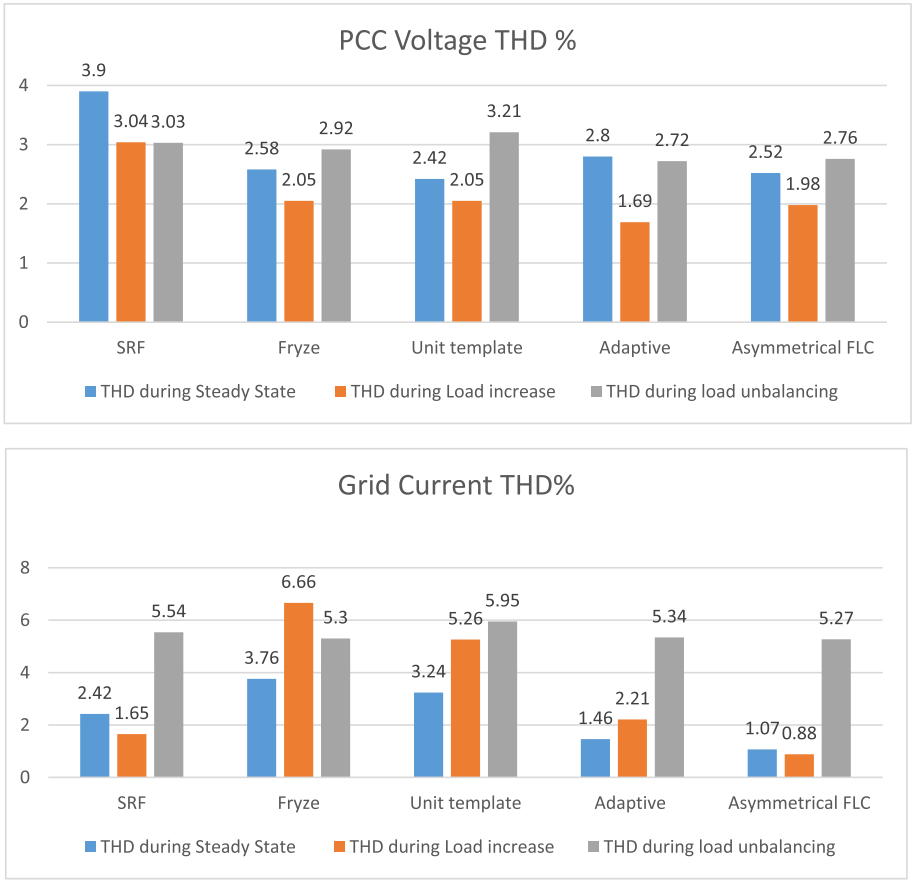


Fig. 10 – THD % of PCC voltage and grid current for different algorithms. Under three conditions (i) Steady-state, (ii) during load increase and (iii) during load unbalancing.

performance when compared with other algorithms on the basis of THD verifies its superiority.

The transient response of DC link voltage control for asymmetrical FLC and other algorithms, i.e. SRF, unit template, conductance fryze and adaptive algorithms are shown in Fig. 11. The response of all the algorithms has been then compared and analyzed for initial interconnection and load perturbation. The load connected to the system is increased from 25 kVA to 50 kVA at 0.3 s. It is observed that the asymmetrical FLC algorithm has the better transient response

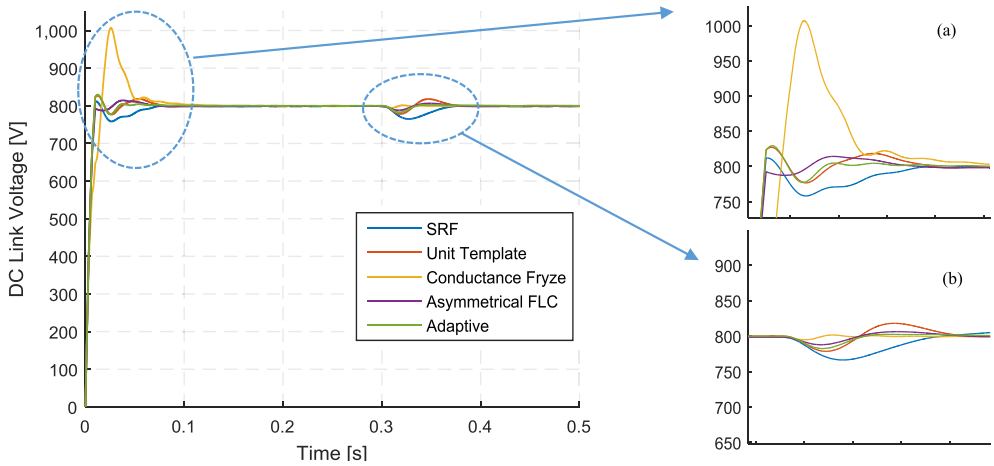


Fig. 11 – Transient performance of the algorithms for both SPV system grid connection and load perturbation conditions. (a) The enlarged transient response of algorithms during SPV system grid connection. (b) The enlarged transient response of algorithms during load perturbation.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

Table 1 – Comparison of the dynamic behavior of proposed asymmetrical FLC algorithm with other algorithms.

Algorithm	During SPV system grid connection			During load perturbation		
	Settling Time (sec)	DC link voltage undershoot (V)	DC link voltage overshoot (V)	Settling Time (sec)	DC link voltage undershoot (V)	DC link voltage overshoot (V)
SRF	0.095	760	812	0.08	760	802
Unit template	0.095	780	810	0.08	780	830
Conductance fryze	0.12	798	1010	0.03	790	810
Adaptive	0.095	785	820	0.03	792	802
Asymmetrical FLC	0.095	792	810	0.03	795	806

during the interconnection of SPV system to the grid as well as during load perturbation transients.

Table 1 shows the comparative analysis of the proposed asymmetrical FLC algorithm with conventional and adaptive control algorithms. Conventional control algorithms i.e. SRF, unit template, and conductance fryze algorithms; and adaptive algorithm are characterized by large undershoot, overshoot and settling time in both SPV system grid connection and load perturbation conditions. The asymmetrical FLC algorithm behaves smoother with less settling time, undershoot and overshoots in both the conditions. Thus, the response of asymmetrical FLC is much superior when SPV system is coupled with grid and when there is load perturbation.

Conclusion

In this paper, an asymmetrical FLC algorithm for PV inverter control has been proposed for SPV system connected to the grid. With the proposed algorithm, the author(s) have maintained the THD level by incorporating fine-tuned functions, and as a result achieved the power quality standards with linear, non-linear and unbalance loads for grid current and PCC voltage. Further, a comparative analysis of control algorithms shows that asymmetrical FLC algorithm has a better transient response over other algorithms under SPV system grid connection as well as load disturbance condition. Simulation investigations have validated the effectiveness of asymmetrical FLC algorithm for SPV system connected to the grid under steady-state as well as dynamic operating conditions. Asymmetrical FLC has been found useful in several applications such as active filtering, VAR generation, and load balancing. The study shall be helpful for power system design engineers working in the area of DERs.

Appendix A

- *Data for SPV Sunmodule SW-230*:

Peak power of array = 10 kW, Short circuit current of module = 8.72 A, open circuit voltage of module = 36.7 V, Cells per module = 60, series resistance = 0.325 Ω, parallel resistance = 50 Ω, Module current at MPP = 7.74 A, Module voltage at MPP = 29.7 V. Solar irradiation = 1000 W/m^2 and temperature = 278 K. Diode ideality factor = 0.9616.

- *Data for DC–DC boost converter*:

D = 0.531–0.493, L = 8.31 mH, C = 255 μF, f_s = 10 kHz.

- *Data for different loads*:
  - Linear load- RL load of 25 kVA with 0.6 lagging pf.
  - Non-linear load- Three-phase bridge rectifier with 10 Ω, 0.8 mH RL load.

REFERENCES

[1] Lasseter R. Microgrids. In: IEEE power engineering society winter meeting; 2002. p. 305–8.
[2] Ellabban O, Abu-Rub H, Blaabjerg F. Renewable energy resources: current status, future prospects and their enabling technology. Renew Sustain Energy Rev 2014. http://dx.doi.org/10.1016/j.rser.2014.07.113.
[3] Dincer I, Acar C. Review and evaluation of hydrogen production methods for better sustainability. Int J Hydrogen Energy September 2015;40(34):11094–111. http://dx.doi.org/10.1016/j.ijhydene.2014.12.035.
[4] Gupta N, Garg R, Kumar P. Sensitivity and reliability models of a PV system connected to grid. Renew Sustain Energy Rev March 2017;69:188–96. http://dx.doi.org/10.1016/j.rser.2016.11.031.
[5] Garg R, Singh A, Gupta S. PV cell models and dynamic simulation of MPPT trackers in MATLAB. Int Conf Comput Sustain Glob Dev (INDIACom) March 2014. http://dx.doi.org/10.1109/IndiaCom.2014.6828003.
[6] Deveci O, Kasnakoglu C. Performance improvement of a photovoltaic system using a controller redesign based on numerical modelling. Int J Hydrogen Energy 3 August 2016;41(29):12634–49. http://dx.doi.org/10.1016/j.ijhydene.2016.05.149.
[7] Amrouche B, Belhaouas N, Achachera S, Boukadoum MT. A software tool for testing maximum power point tracking algorithms using real working conditions. Int J Hydrogen Energy 3 August 2016;41(29):12525–34. http://dx.doi.org/10.1016/j.ijhydene.2016.04.147.
[8] Kaouane M, Boukhelifa A, Cheriti A. Regulated output voltage double switch buck-boost converter for photovoltaic energy application. Int J Hydrogen Energy 7 December 2016;41(45):20847–57. http://dx.doi.org/10.1016/j.ijhydene.2016.06.140.
[9] Ramadan HS. Optimal fractional order PI control applicability for enhanced dynamic behavior of on-grid solar PV systems. Int J Hydrogen Energy 16 February 2017;42(7):4017–31. http://dx.doi.org/10.1016/j.ijhydene.2017.01.122.
[10] Bouzelata Y, Kurt E, Chenni R, Altın N. Design and simulation of a unified power quality conditioner fed by solar energy. Int J Hydrogen Energy 26 November 2015;40(44):15267–77. http://dx.doi.org/10.1016/j.ijhydene.2015.02.077.
[11] Pacas JM, Molina MG, Santos EC. Design of a robust and efficient power electronic interface for the grid integration of solar photovoltaic generation systems. Int J Hydrogen Energy July

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103

2012;37(13):10076–82. http://dx.doi.org/10.1016/j.ijhydene.2011.12.078.
[12] Molina MG, Espejo EJ. Modelling and simulation of grid-connected photovoltaic energy conversion systems. Int J Hydrogen Energy 27 May 2014;39(16):8702–7. http://dx.doi.org/10.1016/j.ijhydene.2013.12.048.
[13] Deveci O, Kasnakoglu C. MIMO nonlinear modelling and robust control of standalone dc photovoltaic systems. Int J Hydrogen Energy April 2017. http://dx.doi.org/10.1016/j.ijhydene.2017.04.033. in press.
[14] Erickson RW. Fundamentals of power electronics. 2nd ed. Secaucus, NJ, USA: Kluwer Academic Publishers; 2000. p. 22–6.
[15] Eshram T, Chapman PL. Comparison of photovoltaic array maximum power point techniques. IEEE Transaction Energy Convers June 2007;22(2):439–49. http://dx.doi.org/10.1109/TEC.2006.874230.
[16] IEEE Standard 1547.2-2008. IEEE Standard for interconnecting distributed resources with electric power systems. 2008.
[17] Koutroulis E, Blaabjerg F. Design optimisation of transformerless grid-connected PV inverters including reliability. IEEE Trans Power Electron 2013. http://dx.doi.org/10.1109/TPEL.2012.2198670.
[18] Ghosh A, Ledwich G. Power quality enhancement using custom power devices. Delhi: Springer International Edition; 2009.
[19] Rub HA, Malinowski M, Al-Hadad K. Power electronics of renewable energy systems, transportation, and industrial application. 1st ed. UK: IEEE Press and Wiley; 2014.
[20] Sebaaly F, Vahedi H, Kanaan HY, Moubayed N, Al-Haddad K. Design and implementation of space vector modulation-based sliding mode control for grid-connected 3L-NPC inverter. IEEE Trans Industrial Electron December 2016;63(12).
[21] Hu J, Shang L, He Y, Zhu ZQ. Direct active and reactive power regulation of grid-connected dc/AC converters using sliding mode control approach. IEEE Trans Power Electron January 2011;26(1).
[22] Gupta N, Garg R, Kumar P. Study of conventional control algorithms for PV-based grid-connected microgrid. In: Proceeding of International conference on intelligent communication, control and devices. Series: Advances in Intelligent systems and computing, vol. 479; 2017, ISBN 978-981-10-1707-0. http://dx.doi.org/10.1007/978-981-10-1708-7_25.
[23] Singh B, Chandra A, Al-Haddad K. Power quality: problems and mitigation techniques. John Wiley & Sons; 2014.
[24] Ghosh A, Joshi A. A new method for load balancing and power factor correction in power distribution system. IEEE Transaction Power Deliv January 2000;15(1):417–22.
[25] Singh B, Kumar S. Modified power balance theory for control of DSTATCOM. In: Proceedings of the joint international conference on power electronics, drives and energy systems (PEDES) & 2010 Power India, New Delhi; Dec. 2010.
[26] Hirofumi A, Edson W, Aredes M. Instantaneous power theory and applications to power conditioning. 1st ed. New Jersey: IEEE Press-Wiley Press Pvt. Ltd; 2007.
[27] Rowey CN, Summersz TJ, Betz RE, Cornforth D. A comparison of instantaneous and fryze power calculations on P-F and Q-V droop in microgrids. In: Proceedings of the 20th Australasian universities power engineering conference, christchurch; Dec. 2010.
[28] Lee CC. Fuzzy logic in control systems: fuzzy logic controller – part I. IEEE Trans Syst Man Cybern April 1990;20:404–18. http://dx.doi.org/10.1109/21.52551.
[29] Chaturvedi DK. Soft computing. 1st ed. Berlin Heidelberg: Springer-Verlag; 2008.
[30] Soufi Y, Bechouat M, Kahla S. Fuzzy-PSO controller design for maximum power point tracking in photovoltaic system. Int J Hydrogen Energy 30 March 2017;42(13):8680–8. http://dx.doi.org/10.1016/j.ijhydene.2016.07.212.
[31] Garcıa P, Torreglosa J, Fernández LM, Jurado F. Optimal energy management system for standalone wind turbine/photovoltaic/hydrogen/battery hybrid system with supervisory control based on fuzzy logic. Int J Hydrogen Energy 4 November 2013;38(33):14146–58. http://dx.doi.org/10.1016/j.ijhydene.2013.08.106.
[32] Bouafia A, Krim F, Gaubert JP. Design and implementation of high performance direct power control of three-phase PWM rectifier, via fuzzy and PI controller for output voltage regulation. Energy Convers Manag 2009;50:6–13.
[33] Cao F, Wang Y. Design of a single-phase grid-connected photovoltaic systems based on fuzzy-PID controller. In: ICIC 2009; 2009. p. 912–9. http://dx.doi.org/10.1007/978-3-642-04020-7_98. LNAI 5755.
[34] Yatak M, Bay O. Fuzzy control of a grid connected three phase two stage photovoltaic system. In: Int conference power engineering, Spain; May 2011. http://dx.doi.org/10.1109/PowerEng.2011.6036477.
[35] Hannan MA, Ghani ZA, Mohamed A, Uddin MN. Real-time testing of a fuzzy-logic-controller-based grid-connected photovoltaic inverter system. IEEE Trans Industry Appl 2015;51(6). http://dx.doi.org/10.1109/TIA.2015.2455025.
[36] Sakhare A, Davari A, Feliachi A. Fuzzy logic control of fuel cell for stand-alone and grid connection. J Power Sources 2004;135:165–76.
[37] Vigneysh T, Kumarappan N. Autonomous operation and control of photovoltaic/solid oxide fuel cell/battery energy storage based microgrid using fuzzy logic controller. Int J Hydrogen Energy 21 January 2016;41(3):1877–91. http://dx.doi.org/10.1016/j.ijhydene.2015.11.022.
[38] Thao NGM, Uchida K. Active and reactive power control techniques based on feedback Linearization and fuzzy logic for three-phase grid-connected photovoltaic inverters. Asian J Control September 2015;17(5):1522–46. http://dx.doi.org/10.1002/asjc.998.
[39] Thao NGM, Uchida K. Control the photovoltaic grid-connected system using fuzzy logic and backstepping approach. In: Proceeding of Asian control conference 2013; June 2013. p. 1–8.
[40] Thao NGM, Uchida K. A two-level control strategy with fuzzy logic for large-scale photovoltaic farms to support grid frequency regulation. Control Eng Pract February 2017;59:77–99. http://dx.doi.org/10.1016/j.conengprac.2016.11.006.
[41] Gupta N, Garg R, Kumar P. Asymmetrical fuzzy logic control to PV module connected micro-grid. In: IEEE India international conference JMI; 2015. http://dx.doi.org/10.1109/INDICON.2015.7443356. India.
[42] Suryanarayana H, Mishra MK. Fuzzy logic based supervision of dc link PI control in a DSTATCOM. In: Annual IEEE India conference, 2008; Dec. 2008. p. 11–3. http://dx.doi.org/10.1109/INDCON.2008.4768766. INDICON.
[43] Kumar P, Mahajan A. Soft computing techniques for the control of an active power filter. IEEE Trans Power Deliv January 2009;24(1).
[44] IEEE Standard 519-2014. IEEE recommended practices and requirements for harmonics control in electric power systems. 2014.

Please cite this article in press as: Gupta N, Garg R, Tuning of asymmetrical fuzzy logic control algorithm for SPV system connected to grid, International Journal of Hydrogen Energy (2017), http://dx.doi.org/10.1016/j.ijhydene.2017.05.103