SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(49th Edition)

JANUARY 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 **Forty ninth** 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 **January 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. Abnormal temperature dependent luminescence behavior of CaSrSiO4:Eu2+ phosphors synthesized via sol-gel strategy, Hyun-Joo Woo, ***3.M. Jayasimhadri***, and Kiwan Jang, Applied Physics, DTU

2. An empirical study for software change prediction using imbalanced data, ***3.Ruchika Malhotra*** & ***6.1.Megha Khanna***, software engineering, DTU

3. An Energy Aware Cost Effective Scheduling Framework for Heterogeneous Cluster System, ***3.Neetesh Kumar*** and Deo Prakash Vidyarthi, CSE, DTU

4. An Object-Oriented MDP Representation for Robotic planning, ***8.Aasheesh Singh,*** Electronics & Communication, DTU

5. Commuters' Exposure to PM2.5: A Case Study of Delhi, ***3.Rajeev Kumar*** Mishra, ***6.Amrit Kumar*** and Ankita Shukla, Environmental Engineering, DTU

6. Comparison of Deflection and slope of Cantilever Beam with Analytical and Finite Element Method for Different Loading Conditions, ***3.Sanjay Kumar***, DTU

7. Design of Solar Powered UAV, ***3.Janardan Prasad Kesari*** and ***7.Abhishek Shakya***, Mechanical Engineering, DTU

8. Development and electrochemical performances of Li3V2(PO4)3 and Li4Ti5O12 materials for lithium-ion battery***, 6.Rakesh Saroha, 3.Amrish K. Panwar, 6.Aditya Jain, Jitendra Singh & 6.SandeepVerma,*** Applied Physics, DTU

9. Estimation of intrinsic work function of multilayer graphene by probing with electrostatic force microscopy, ***6.1.Anshika Singh***, Puspendu Guha,***3.Amrish K. Panwar***, ***3.Pawan K. Tyagi***, Applied Physics, DTU

10. Investigation of Charge-Separation/Change in Dipole Moment of 7- Azaindole: Quantitative measurement using Solvatochromic shifts and Computational approaches, ***3.Mohan Singh Mehata,*** Ajay K. Singh and ***3.Ravindra Kumar Sinha,*** Applied Physics, DTU

11. Localized & self adaptive audio watermarking algorithm in the wavelet domain, Arashdeep Kaur, Malay Kishore Dutta, K.M. Soni and ***3.Nidhi Taneja,*** DTU

12. Mechanical properties and microstructure studies in Friction Stir Welding (FSW) joints of dissimilar alloy – a review , ***6.H. Mehdi and 3.R.S.Mishra***, Mechanical, DTU

13. Mechanisms of plasma-assisted catalyzed growth of carbon nanofibres: a theoretical modeling, ***3.R Gupta, 3.S C Sharma and 3.R Sharma***, Applied Physics, DTU

14. Methods for the Screening and Treatment of Spinocerebellar Ataxias, ***6.Pushpendra Mani Mishra*** and ***6.Dr. Vimal Kishor Singh***, Biotechnology, DTU

15. Minimally coupled scalar field cosmology in anisotropic cosmological model, ***3.C P SINGH*** and ***6.MILAN SRIVASTAVA,*** Applied Mathematics, DTU

16. Numerical Simulation of NACA 2415 Airfoil at Different Low Reynolds Numbers and Angles of Attack, ***7.1.Aakash Srivastava, 7.1Vatsalya Tiwari*** and ***3.Raj Kumar Singh,*** Mechanical Engineering, DTU

17. On the evolution and electronic properties of self-assembled gold nanowires, ***3.Neha Tyagi,*** Neeraj K. Jaiswal, Goran Kovacˇevic and Pankaj Srivastava, Applied Physics, DTU

18. Optimization of Synthesis Technique and Luminescent Properties in Eu3+-activated NaCaPO4 Phosphor for Solid State Lighting Applications, B.V. Ratnam***, 6.Mukesh K. Sahu***, ***6.Amit K. Vishwakarma***, ***6.Kaushal Jha***, Hyun-Joo Woo, Kiwan Jang and ***3.M. Jayasimhadri***, Applied Physics, DTU

19. Pulse TIG welding: Process, Automation and Control, ***3.P.K. Baghel*** and ***3.D.S. Nagesh,*** Mechanical Engineering, DTU

20. QFD Concepts Integrated with Design of Indian Military Vehicles, MK Shukla, BS Chauhan and ***6.1Ashutosh Kumar Rai***, Mechanical Engineering, DTU

21. Quantum analysis based extraction of frequency dependent intrinsic and extrinsic parameters for GEWE-SiNW MOSFET, ***6.Neha Gupta*** and ***6.Rishu Chaujar,*** Physics, DTU

22. Stress Analysis at Contact Region of Rail-Wheel: Review, ***3.Roop Lal, 3.R.C. Singh*** and ***7.1.Davendra Singh***, DTU

23. Study of corrosive fatigue and life enhancement of low pressure steam turbine blade using friction dampers, Loveleen Kumar Bhagi, ***3.Vikas Rastogi*** and Pardeep Gupta, Mechanical, Production, Industrial and Automobile, DTU

24. Synthesis and optical properties of core-multi-shell CdSe/CdS/ZnS quantum dots: Surface modifications, ***3.R.K. Ratnesh*** and ***3.Mohan Singh Mehata***, Applied Physics, DTU

25. Thermodynamic Evaluation of PTC based Organic Rankine Cycle for Power & Cooling, ***3.Devendra Kumar Gupta***, ***3.Rajesh Kumar*** and ***3.Naveen Kumar***, Mechanical, 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*

Accepted Manuscript



Abnormal temperature dependent luminescence behavior of $CaSrSiO_4:Eu^{2+}$ phosphors synthesized via sol-gel strategy

Hyun-Joo Woo, M. Jayasimhadri, Kiwan Jang

PII: S0925-8388(17)30200-1

DOI: 10.1016/j.jallcom.2017.01.184

Reference: JALCOM 40551

To appear in: *Journal of Alloys and Compounds*

Received Date: 24 November 2016

Revised Date: 11 January 2017

Accepted Date: 18 January 2017

Please cite this article as: H.-J. Woo, M. Jayasimhadri, K. Jang, Abnormal temperature dependent luminescence behavior of $CaSrSiO_4:Eu^{2+}$ phosphors synthesized via sol-gel strategy, *Journal of Alloys and Compounds* (2017), doi: 10.1016/j.jallcom.2017.01.184.

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

Abnormal Temperature dependent luminescence behavior of $CaSrSiO_4:Eu^{2+}$ phosphors synthesized via sol-gel strategy

Hyun-Joo Woo,[a] M. Jayasimhadri,[b] and Kiwan Jang[c] *

[a]*Research Institute of Basic Sciences, Changwon Nat'l University, Changwon 51140, Republic of Korea*

[b]*Department of Applied Physics, Delhi Technological University, Delhi -110042, India*

[c]*Department of Physics, Changwon Nat'l University, Changwon 51140, Republic of Korea*

ACCEPTED MANUSCRIPT

*Corresponding Author

Prof. Kiwan Jang

Department of Physics

Changwon National University

9, Sarim-dong, Changwon, GN, 51140

Republic of Korea

Email: kwjang@changwon.ac.kr

Ph. No. +82-055-213-3427, Fax: +82-55-213-0264

Abstract

A silicate based crystal doped with Eu^{2+} phosphors have been synthesized by sol-gel method using propylene glycol modified silane (PGMS). The structural and luminescence properties of the as-prepared samples were well characterized by using x-ray diffractometer (XRD) & photoluminescence spectroscopy. Thermal stability of the prepared phosphors was also measured and analyzed in detail. The results revealed a new fact that longer wavelength emission band position is relatively more sensitive to the sample temperature for higher dopant concentrations when compared to the other emission bands. Such observations were not found in the samples prepared by solid phase reaction method. The $CaSrSiO_4$ phosphor doped with 3 mol % of Eu^{2+} showed unique optical properties that consist of a shift in emission spectrum towards shorter wavelength along with a decreased band width (FWHM). The emission color tunablity with the increase in temperature and concentration of the $CaSrSiO_4$ phosphor were also discussed by CIE chromaticity diagram.

Keywords: $CaSrSiO_4$; Photoluminescence; Thermal Stability; pc-white LED Phosphor; PGMS.

ACCEPTED MANUSCRIPT

1. Introduction

White light emitting diodes (w-LEDs) are considered as new lighting sources for the present and next generations due to their unique properties such as high efficiency, long lifetime and environmental friendliness[1-8]. In general, w-LEDs can be realized by three approaches. Those are (i) mixing of red, green and blue LEDs (ii) UV emitting LED combined with red, green and blue phosphors and (iii) blue LED combined with yellow phosphor. Among them, techniques (ii) and (iii) are referred as phosphor converted white LED (pc-white LEDs). The pc-white LEDs are a preferential approach compared to the mixing of red, green and blue LEDs because of their simple structure and low cost etc.,. The demand for the high power pc-LEDs is increasing day by day in various applications such as traffic signal lighting, sport complex lighting, etc.,. The phosphors used in pc- white LEDs play a key role to obtain a lighting system along with high thermal stability and quality output[7-12]. Thus studying the thermal stability of phosphor becomes very important to develop new kinds of phosphor materials. Nowadays, the pc-white LEDs have been developed with higher power where the chips release high temperatures (e.g., in the range of 100 °C-150 °C). The temperature is one out of several factors that determine the quality of LEDs such as color coordinates, etc.,. So, understanding the luminescence behavior against temperature is important and necessary to improve the optical performance of pc-white LEDs.

In this article, we report the temperature dependent luminescence behavior of Eu^{2+} doped $CaSrSiO_4$ prepared through sol-gel process by using PGMS as silicate source for the first time. The significant result obtained for the first time in this manuscript that longer wavelength emission band FWHM and intensities are relatively more sensitive to the sample temperature for higher dopant concentrations. This study would help to understand the mechanisms responsible for the luminescence quenching, variation in chromaticity coordinates at high temperatures and for developing new materials for pc-white LEDs.

2. Materials and methods

2.1. Materials

Strontium nitrate ($Sr(NO_3)_2$), calcium nitrate hexahydrate ($Ca(NO_3)_6{\cdot}6H_2O$), europium nitrate hexahydrate ($Eu(NO_3)_3{\cdot}6H_2O$) and tetraethyl orthosilicate (TEOS) were purchased from sigma–Aldrich. Citric acid, Propylene glycol and Hydrochloric acid (HCl) were received from Daejung chemicals. All the above mentioned chemicals were used as received.

2.2. Preparation of PGMS solution

Initially PGMS solution was prepared as per our previous report [13]. Briefly, 22.3 mL of TEOS and 22.3 mL of Propylene glycol were mixed together in a round bottomed flask and stirred for 48 h at 80 °C to get a turbid bilayered liquid. Exactly, 15 mL of distilled water was mixed with 5 mL of turbid bilayered liquid (TEOS + propylene glycol) solution. To the above mixture, 0.1 mL of concentrated hydrochloric acid was added slowly and stirred well to get a transparent PGMS solution. This solution was used for the further preparation of the $CaSrSiO_4:Eu^{2+}$ phosphor.

2.3. Preparation of $CaSrSiO_4:Eu^{2+}$ phosphor

The $Ca_{1-x}Sr_{1-x}SiO_4$: $2xEu^{2+}$ ($2x$ = 0.0025, and 0.03) phosphors were synthesized via sol-gel strategy. The prepared PGMS solution was taken in a beaker along with the aqueous solution containing the mixture of $Sr(NO_3)_2$, $Ca(No_3)_2{\cdot}6H_2O$, $Eu(No_3)_3{\cdot}6H_2O$ and citric acid. This mixture was initially stirred at 80 °C for 2 h to get homogeneous solution. The propylene glycol was added to it under stirring for another 24 h at 120 °C to obtain the polymerized gel. It was then subjected to thermal treatment in ambient atmosphere at 1,300 °C for 6 h to decompose the organic compound in the polymerized gel sample. Thus, a pure white $Ca_{1-x}Sr_{1-x}SiO_4:2xEu^{3+}$ phosphor powder was obtained. These phosphors were then heat-treated at

1,100 °C under reduced atmosphere (5% H_2 +95% N_2) for 4 h in order to reduce Eu^{3+} into Eu^{2+} in $CaSrSiO_4$ host. The obtained products were finely ground and characterized.

2.4. Characterization techniques

The X–ray diffraction studies (XRD, X'Pert Pro, PANalytical) were investigated to understand the change in crystalline features of the samples due to the introduction of different Eu^{2+} concentrations and temperature. The finely ground samples were mounted onto the polymeric sample holder. It was then analyzed between the range of 2θ from 10 to 80° with the scan speed of 0.02°/sec using $CuK_{\alpha 1}$ radiation (λ=1.5406 Å) at 40 kV and 30 mA. To study the luminescent features of Eu^{2+} doped in $CaSrSiO_4$, the photoluminescence excitation and emission spectra of all the samples were recorded using a fluorescent spectrofluorometer (JASCO FP-8500, Japan) equipped with an integrating sphere (ISF-834). The samples were mounted on an aluminum sample holder where the Xe lamp was used as an excitation light source. The thermal stability for the studied samples was also measured by using a homemade heating unit and an ocean optics spectrophotometer (HR4000). It was performed by placing the powder samples in a copper sample holder which is connected with a temperature controller and exciting the samples with 365 nm ultraviolet from conventional UV lamp. The emission intensity was measured at every 10 °C during an increase from 20 to 200 °C and then decrease from 200 to 20 °C.

3. RESULTS AND DISCUSSION

3.1. Effect of dopant Concentration and temperature on structural properties

The present study mainly focused to reveal the effect of dopant concentration ($2xEu^{2+}$) and temperature on the $Ca_{1-X}Sr_{1-X}SiO_4$ crystal structure. The variation in crystal phase of $Ca_{1-X}Sr_{1-X}SiO_4$:$2xEu^{2+}$ with the temperature were also analyzed within the range from 20 °C to 200 °C. Figure 1(a) shows the XRD patterns of $Ca_{1-X}Sr_{1-X}SiO_4$:$2xEu^{2+}$ while varying the

dopant concentrations (2x = 0.0025 and 0.03). The XRD patterns of all the phosphors are well matched with the standard ICSD card (49006) without any changes in the crystal phase regardless of Eu^{2+} concentration. This confirms that the Eu^{2+} ions doped in the samples did not affect the crystal phase of $Ca_{1-X}Sr_{1-X}SiO_4$. Based on the XRD patterns observed for all these sol-gel synthesized phosphors, it is to be confirmed that an orthorhombic α′-$CaSrSiO_4$ phase along with the space group of Pmnb. Moreover the lattice parameters are found to be a=5.6049, b=6.9623 and c=9.483Å. The temperature dependent XRD patterns of samples with two dopant concentrations, $Ca_{0.99875}Sr_{0.99875}SiO_4$:$0.0025Eu^{2+}$ & $Ca_{0.985}Sr_{0.985}SiO_4$:$0.03Eu^{2+}$, are shown in Figure 1 (b) and (c). The observation of all the peaks in the XRD patterns clearly confirms that the increase in temperature up to 200 °C did not affect the $CaSrSiO_4$ crystal phase.

3.2. photoluminescence (PL) studies

Figure 2 illustrates the excitation and emission spectra of $Ca_{0.99875}Sr_{0.99875}SiO_4$: $0.0025Eu^{2+}$ & $Ca_{0.985}Sr_{0.985}SiO_4$:$0.03Eu^{2+}$ phosphors prepared by sol-gel method using PGMS. The excitation spectra of both samples were measured in a very wide range of 200 to 550 nm. This kind of broad excitation bands can commonly be ascribed due to the splitting of $4f^65d^1$ level of Eu^{2+} ions, which have strong interaction with neighboring anion ligands in the $CaSrSiO_4$ crystal lattice. In addition, it was observed that the emission spectrum of the sample doped with 0.03 Eu^{2+} shifted to the longer wavelength region and broadened when compared to the sample doped with 0.0025 Eu^{2+}. The broad emission observed in the visible region might have occurred from the $4f^65d^1 \rightarrow 4f^7$ ($^8S_{7/2}$) transition of Eu^{2+} ions doped in $CaSrSiO_4$. A red shift occurred with increasing Eu^{2+} concentration could be clearly observed, and this red shift was much greater when compared with the conventionally synthesized $CaSrSiO_4$:Eu^{2+} through various methods including solvo-thermal and solid-state techniques[14-20]. The luminescent characteristics of the observed spectra clearly exhibit that the intensity and band position of emission and excitation spectra are strongly dependent on the dopant concentration (Eu^{2+}).

Usually, higher concentration of Eu^{2+} ions increases the energy transfer between dopant ions,

thus affecting the spectral properties. In this case, the shape of spectral line are continuously changed and shifted in longer wavelength. However, in the present work, the observed spectral shape was suddenly changed, which may be due to the increase of occupation of Eu^{2+} in the particular site above the certain Eu^{2+} concentration. The authors thus assume that the change of spectral shape do not originate from the energy transfer between dopant ions but occupation of Eu^{2+} in the certain site.

In general, the emission intensity of any emission band usually depends upon the intensity of excitation band. However, the emission intensity is higher for lower intensity excitation band and vice-versa in the present work, which is unique. To reveal this phenomenon, we measured the quantum efficiencies for the samples doped with 0.0025 Eu^{2+} and 0.03 Eu^{2+}. The quantum efficiency of the sample shown the upper excitation density was 35%, on the other hand, the sample shown for the lower excitation density is 67%. The authors assume that this unique property has been observed due to the difference in quantum efficiency.

3.3. Temperature dependent luminescence behavior

In pc-white LED applications, an important property is the thermal stability of colors that gets emitted from the phosphor. Figure 3 represents the temperature dependent emission spectra of $Ca_{0.99875}Sr_{0.99875}SiO_4$:0.0025$Eu^{2+}$ & $Ca_{0.985}Sr_{0.985}SiO_4$:0.03$Eu^{2+}$ phosphors under the excitation of UV lamp with 365 nm. The spectra show that the decreased emission intensities against increasing temperature which is due to the temperature quenching process. The emission intensity of $Ca_{0.99875}Sr_{0.99875}SiO_4$:0.0025$Eu^{2+}$ in figure 3 (a) shows 13% decrease in the efficiency at the peak luminescence position at 150 °C which is a maximum junction temperature of pc-LEDs [21]. Furthermore, decrease in the overall luminous intensity with increase in temperature was observed but no changes have been observed in the FWHM and shift in the luminescence peak position from Figure 3 (a). Figure 3(b) shows a red shift in the emission spectrum as the concentration of Eu^{2+} increased to 0.03. Also, the luminous intensity was decreased by 26% at 150 °C which is almost twice the magnitude of intensity of previous one. However, the FWHM in this sample was narrowed by 23 nm at

200 °C and also the luminescence peak position shifted from 580 nm to 520 nm. The normalized emission spectra of 3 (a & b) are shown in figures 3 (c) and (d), respectively, which clearly exhibit the variation in FWHM and peak position for 2x= 0.03 concentration. This is a peculiar property found in the samples synthesized through sol-gel strategy for high concentrations using PGMS as silicate source.

Each emission spectrum of $Ca_{1-x}Sr_{1-x}SiO_4:2xEu^{2+}$ has been divided into three different possible emission bands using Gaussian fitting and demonstrated as shown in Fig. 4 (a) to (d). In general, Eu^{2+} ions substituted into two different sites in M_2SiO_4(M: Ba, Ca, Sr), namely, Eu(I) and Eu(II) sites. Hence, PL spectra of Eu^{2+} consist of two emission bands: Eu(I) of shorter wavelength emission and Eu(II) of longer wavelength. In the case of $Ca_2SiO_4:Eu^{2+}$ and $Ba_2SiO_4:Eu^{2+}$, they are overlapped and appear as a single band, which is a non-Gaussian form. It is also known that there are two Ca/Sr sites in CaSrSiO4 structure with different coordination numbers, 10-coordinated Ca1/Sr1 site and 9 coordinated Ca2/Sr2 site and each of them is disordered by Ca/Sr with occupancy 1/2[22]. Thus, it is expected that the PL spectra of Eu^{2+} consist of two emission bands. However, in the present work, the PL spectra of Eu^{2+} consist of three emission bands, which are called as band A, B and C in this paper.

A further analysis on the temperature dependent luminescence intensity of individual splits shows clearly that the B and C are highly dependent on the sample temperature. The ionic radius of Eu^{2+} (1.35Å for coordination number CN=10 and 1.25Å for CN=8) is larger than that of Ca^{2+} ions (1.23 Å for CN=10 and 1.12 Å for CN 8), and approximately equal or smaller than the Sr^{2+} ion (1.36 Å for CN=10 and 1.26 Å for CN=8) [23]. So, the Sr^{2+} sites are the most favorable site for replacement of Eu^{2+} ions when compared to Ca^{2+} sites in the host lattice. Based on this, it is considered that most of the Eu^{2+} ions are expected to occupy the most favorable Sr^{2+} site and then a few at the less favorable Ca^{2+} site at low Eu^{2+}

concentrations. The luminescence intensity of the band C increases with Eu^{2+} concentration, which may be due to the Eu^{2+} ions replaced at the Ca^{2+} site.

An increase in Eu^{2+} concentrations promotes the more occupancy of Eu^{2+} in Ca^{2+} site. Thus, the band A shown in Figure 4 (a–d) may be ascribed to the Eu^{2+} ions in Sr^{2+} sites, while the bands B and C arise from the Eu^{2+} ions in Ca^{2+} sites. So, the relative intensity between band A (Sr^{2+} site) and band C (Ca^{2+} site) increases with the Eu^{2+} concentration due to enhancement in band C. Moreover, the occupancy of larger ionic size Eu^{2+} ions in smaller Ca^{2+} sites increases the crystal field strength of host lattice due to decrease in bond length. This ultimately led to an overall red shift of the photoluminescence. This analysis agrees with the experimental results as shown in Figures 4 (a) to (d).

In Figure 4 (e-f), it is clearly shown that the trend of emission intensity of band C highly depends on the sample temperature as compared to the band A and B. When the temperature of the phosphors doped with 2x=0.0025 Eu^{2+} was increased from 20 °C to 200 °C, the emission intensity due to the band B was increased a little, while those due to the band A and C were found to decrease as much as that due to the band B was increased. This result may explain the trend of the emission spectrum, which showed an overall decrease without any decrease in the FHWM of the wavelength or shift in the luminescence peak wavelength. On the other hand, when the concentration of Eu^{2+} was increased to 0.03, the emission intensity due to band A was little bit decreased, which had been expected that Eu^{2+} ions might occupy at Sr^{2+} site. The emission intensity of the band C due to Eu^{2+} ions substituted in the Ca site was more significantly decreased than band A and B. This phenomenon may significantly effect on the overall emission spectrum. However, the order of the emission intensity band C > band B > band A of the 2x= 0.03 Eu^{2+} sample at 20 °C was modified into band B > band A > band C at 200 °C, which is not observed for low dopant concentration (2x= 0.0025). Based on these results, it was ensured that the temperature highly influence on each emission band of a spectrum of single phosphor at high concentrations. It

could be clearly understood that the emission intensity of band C was significantly decreased as the temperature of the sample increases, which also resulted in the decreasing FWHM and shift in the emission spectrum towards shorter wavelength. In general, the FWHM of a band increases with temperature due to the phonon contribution to emission spectrum. However our results showed unusual phenomenon that the FWHM decreases with temperature due to the difference in thermal stability of each emission band.

The observed blue shifts in emission may be ascribed to two mechanisms, one is the tunneling from excited states of the low-energy emission band to excited states of the high-energy emission band by assistance of thermally active phonons and another is the lower energy site is quenched faster than the higher energy one, which is thermally related energy back transfer from Eu^{2+} ions to the host. In our opinion, the former mechanism is dominant in the phosphor doped with $2x=0.0025$ Eu^{2+}, while the latter is dominant in the sample doped with $2x=0.03$ Eu^{2+}. This can explain the reduction in the FWHM and a blue shifting of emission which results from the change of intensity ratio between two emission bands.

In order to know the effect temperature and dopant concentration on emission color of the synthesized phosphor, the calculated CIE coordinates based on the emission spectra for four samples (M1-M4) are indicated in the CIE chromaticity diagram (Figure 5). It has been observed that both the temperature and dopant concentration significantly affected the emission color of the phosphor, which is varying from green to yellow region.

4. CONCLUSIONS

In summary, the emission spectra of the $Ca_{1-X}Sr_{1-X}SiO_4:2xEu^{2+}$ phosphors prepared by the Sol-gel method using PGMS showed a red shift with increasing Eu^{2+} concentrations. The temperature dependent photoluminescence studies show a decrease in emission intensity of both $2x= 0.0025$ and 0.03 Eu^{2+} doped phosphors with increasing temperatures. However, the $2x= 0.0025$ Eu^{2+} doped phosphor showed a decrease in the overall emission efficiency

without making any changes in FWHM and emission wavelength. But, the phosphor with 2x= 0.03 Eu^{2+} showed a decrease in FWHM along with a shift in the emission peak wavelength to a shorter wavelength. This is the first of its kind to report the site specific temperature dependent luminescence properties of a bulk phosphor as the longer wavelength emission band position and FWHM is relatively more sensitive to the sample temperature for higher dopant concentrations. The CIE coordinates revealed the effect of temperature on the variation of color of synthesized phosphor, which will be more suitable for solid state lighting applications.

Acknowledgments

This work was supported by the Priority Research Centers Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (NRF-2016R1A6A3A01007064) and was also partially supported by the Korean government (NRF–2015–056886).

References

[1] C.C. Lin and R. S. Liu, Advances in phosphors for light-emitting diodes, J. Phys. Chem. Lett. 2 (2011) 1268-1277.

[2] S. Ye, F. Xiao, Y. X. Pan, Y. Y. Ma and Q. Y. Zhang, Phosphors in phosphor-converted white light-emitting diodes: Recent advances in materials, techniques and properties, Mater. Sci. Eng. R. 71 (2010) 1-34.

[3] D. A. Steigerwald, J. C. Bhat, D. Collins, R. M. Fletcher, M. O. Holcomb, M. J. Ludowise, P. S. Martin, S. L. Rudaz, Illumination with solid state lighting technology, IEEE J. Sel. Quant. 8 (2002) 310-320.

[4] MM. R. Krames, O. B. Shechekin, R. Mueller-Mach, G. O. Mueller, L. Zhou, G. Harber, M. G. Craford, Status and Future of High-Power Light-Emitting Diodes for Solid-State Lighting, J. Disp. Technol. 3 (2007) 160-175.

[5] M. Yamada, Y. Narukawa, H. Tamaki, Y. Mukai, T. Mukai, A Methodological Study of the Best Solution for Generating White Light Using Nitride-Based Ligh-Emitting Diodes, IEICE. Transantions. Electonics. E88-C (2005) 1860-1871.

[6] Y. Narukawa, White-Light LEDs, Opt. Photon. New. 15 (2004) 24-29.

[7] J. K. Sheu, S. J. Chang, C. H. kuo, Y. K. Su, L. W. Wu, Y. C. Lin, W. C. Lai, J. M. Tsai, G. C. Chi, R. K. Wu, White-Light Emission From Near UV InGaN-GaN LED Chip Precoated With Blue/Green/Red phosphors, IEEE Photonic Tech. Lett. 15 (2003) 18-20.

[8] M. S. Shur and A. Zukauskas, Solid-State Lighting: Toward Superior Illumination, Proc. IEEE. 93 (2005) 1691-1703.

[9] K. N. Shinde, S. J. Dhoble, H. C. Swart, and K. Park, Phosphate Phosphors for Solid–State Lighting, Springer-Verlag, Berlin, 2012.

[10] Philippe F. Smet, Anthony B. Parmentier, and Dirk Poelman, Selecting Conversion Phosphors for White Light-Emitting Diodes, J. Electrochemical Society. 158 (2011) R37-R54.

[11] C. Feldman, T. Justel, C. R. Ronda, P. J. Schmidt, Inorganic Luminescent Materials: 100 Years of Research and Application, Adv. Funct. Mate. 13 (2003) 511-516.

[12] L. S. Rohwer, A. M. Stivastava, Development of Phosphors For LEDS, Electrochem. Soc. Interface. 12 (2003) 36-40.

[13] H. J. Woo, S. Gandhi, B. J. Kwon, D. S. Shin, S. S. Yi, J. H. Jeong and K. Jang, Soluble silica assisted synthesis and luminescent characteristics of yellow emitting $CaSrSiO_4:Eu^{2+}$ phosphors for warm white light production, Ceram. Int. 41 (2015) 5547-5553.

[14] S. Gandhi, K. Thandavan, B. J. Kwon, H. J. Woo, K. Jang, D. S. Shin, Solvothermal synthesis of red and green emitting $Ca_{1.65}Sr_{0.35}SiO_4:Eu^{3+}$ and $Ca_{1.65}Sr_{0.35}SiO_4:Eu^{2+}$ phosphors for solid-state lighting applications, Ceram. Int. 40 (2014) 5245-5254.

[15] W. J. Park, Y. H. Song, D. H. Yoon, Synthesis and luminescent characteristics of $Ca_{2-x}Sr_xSiO_4:Eu^{2+}$ as a potential green-emitting phosphor for near UV-white LED applications, Mater. Sci. Eng. B. 173 (2010) 76-79.

[16] W. Zhi Jun, Y. Z. P ing, G. Q. Lin, L. P. Lai, F. G. Sheng, Luminescence characteristics of Eu^{2+} activated Ca_2SiO_4, Sr_2SiO_4 and Ba_2SiO_4 phosphors for white LEDs, Chin. Phys. B. 18 (2009) 2068-2071.

[17] H. Yu, Y. Lai, G. Gao, L. Kong, G. Li, S. Gan, G. Hong, Photoluminescence and energy transfer studies on Eu^{2+} and Ce^{3+} co-doped $SrCaSiO_4$ for white light emitting diodes, J. Alloy. Compd. 509 (2011) 6635-6639.

[18] S. Tezuka, Y. Sato, T. Komukai, Y. Takatsuka, H. Kato, M. Kakihana, Eu^{2+}-activated $CaSrSiO_4$: a new red-emitting oxide phosphor for white light emitting diodes, Appl. Phys. Express. 6 (2013), 072101.

[19] S. Gandhi, K. Thandavan, B.J. Kwon, H.J. Woo, C.H. Kim, S.S. Yi, J.H. Jeong, D.S. Shin, K. Jang, A highly efficient warm white light emitting Eu^{2+}-activated silicate host : another striking application of mesoporous silica, J. Mater. Chem. C. 2 (2014) 6630-6636.

[20] B. J. Kwon, S. Gandhi, H.J. Woo, K. Jang, D.S. Shin, K. Kim, S. Jeong, Optical properties and carrier dynamics of $CaSrSiO_4:Eu^{3+}$ phosphors prepared by using the solid-state reaction method, J. Korean Phys. Soc. 64 (2014) 1721-1725.

[21] Y. Gu and N. Narendran, A non-contact method for determining junction temperature of phosphor -converted white LEDs, Proc. SPIE. 5187 (2004) 107-114.

[22] Shihai Miao, Zhiguo Xia, Maxim S. Molokeev, Mingyue Chen, Jie Zhang, Quanlin Liu, Effectt of Al/Si Substitution on the Structure and Luminescence Properties of $CaSrSiO_4:Ce^{3+}$ Phosphor: Analysis based on the Polyhedral Distortion, J. Mater. Chem. C, 3, 4616-4622(2015).

[23] R. D. Shannon, Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides, Acta Crystallogr. A. A32 (1976) 751-767.

Figure captions:

Fig. 1. (a) X-ray diffraction patterns of as-prepared $Ca_{1-X}Sr_{1-X}SiO_4:Eu^{2+}$ phosphors with different concentration of Eu^{2+}. The figures (b) and (c) show the temperature dependent powder XRD patterns of the $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ & $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ samples, respectively.

Fig. 2. Excitation and emission spectra of the $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ & $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ phosphors.

Fig. 3. Temperature dependent emission spectra of sol-gel synthesized $Ca_{0.99875}$ $Sr_{0.99875}S$ iO_4: 0.0025 Eu^{2+} & $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ phosphors (a and b). The figures c and d show the normalized emission spectra of $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ & $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ phosphors, respectively.

Fig. 4. Gaussian fitting carried out on the photoluminescence of $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ at 20 °C (a) & $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ at 20 °C (b) $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ at 200 °C (c) & $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ at 200 °C (d) phosphors. Temperature dependent photoluminescence intensity of the three emission bands corresponding to (e) 0.0025 and (f) 0.03 Eu^{2+} doped in $CaSrSiO_4$. The three emission bands were mentioned by band A, B and C.

Fig. 5. The CIE chromaticity diagram for $Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ at 20 °C,

$Ca_{0.985}Sr_{0.985}SiO_4:0.03Eu^{2+}$ at 200 °C, $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ at 20 °C and $Ca_{0.99875}Sr_{0.99875}SiO_4:0.0025Eu^{2+}$ at 200 °C indicated as M1, M2, M3 and M4.

Figure 1.

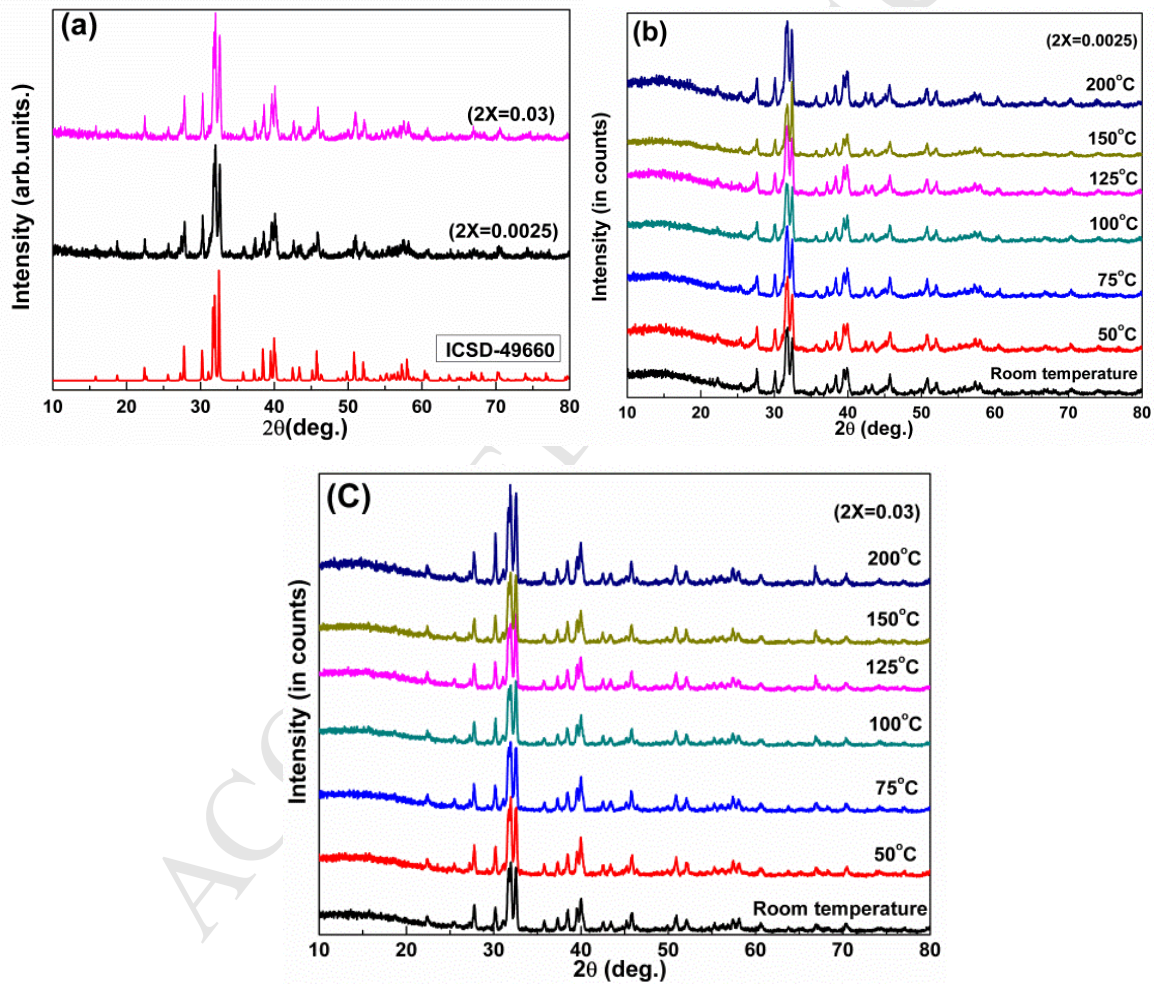


Figure 2.

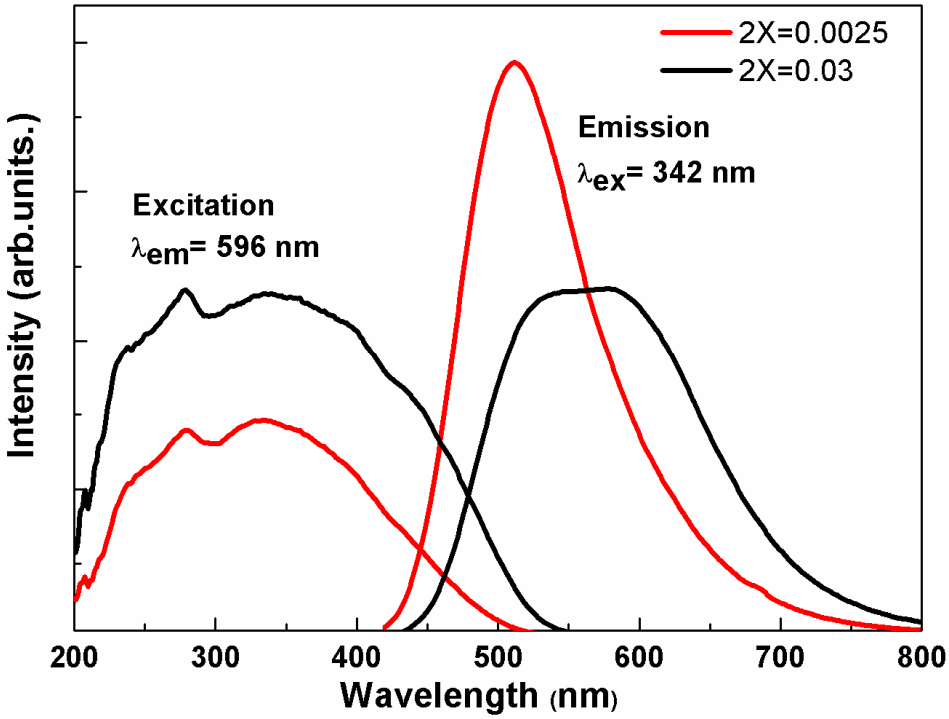

2X=0.0025
2X=0.03
Emission
λex= 342 nm
Excitation
λem= 596 nm
Intensity (arb.units.)
200
300
400
500
600
700
800
Wavelength (nm)

Figure 3.

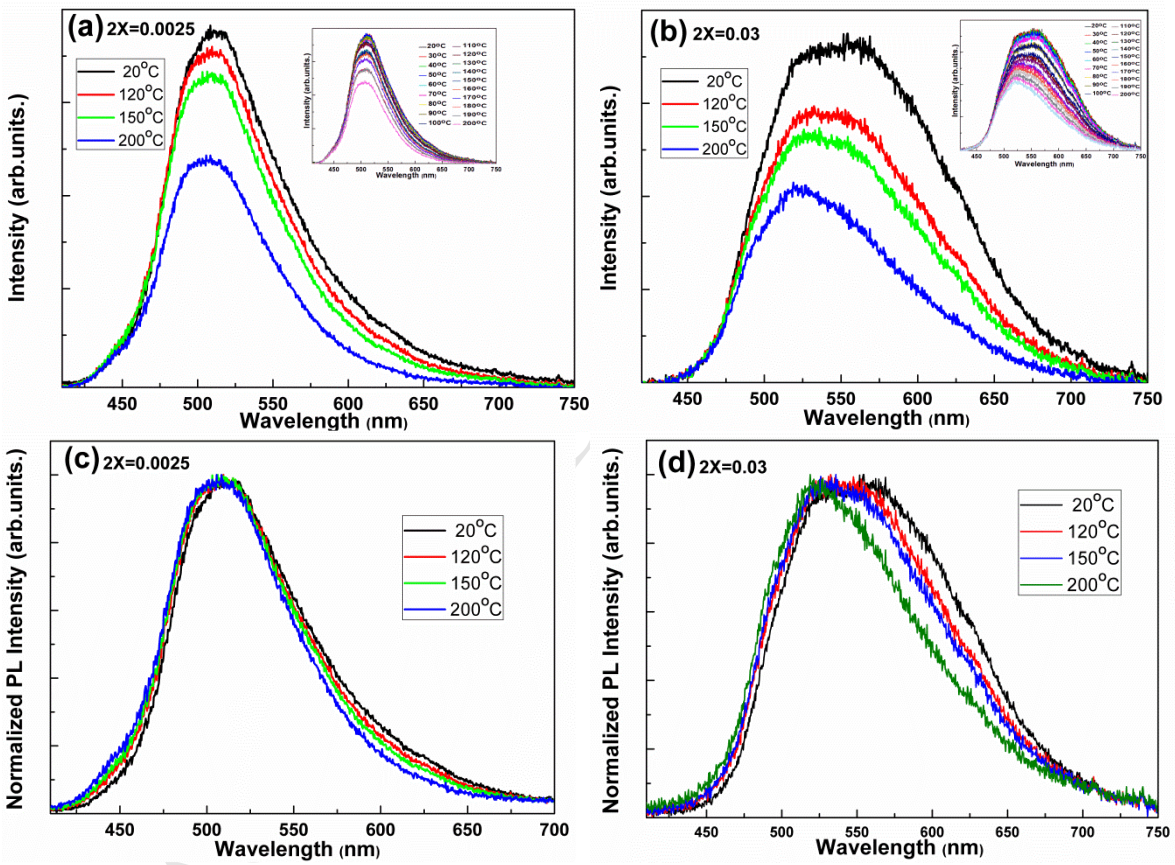
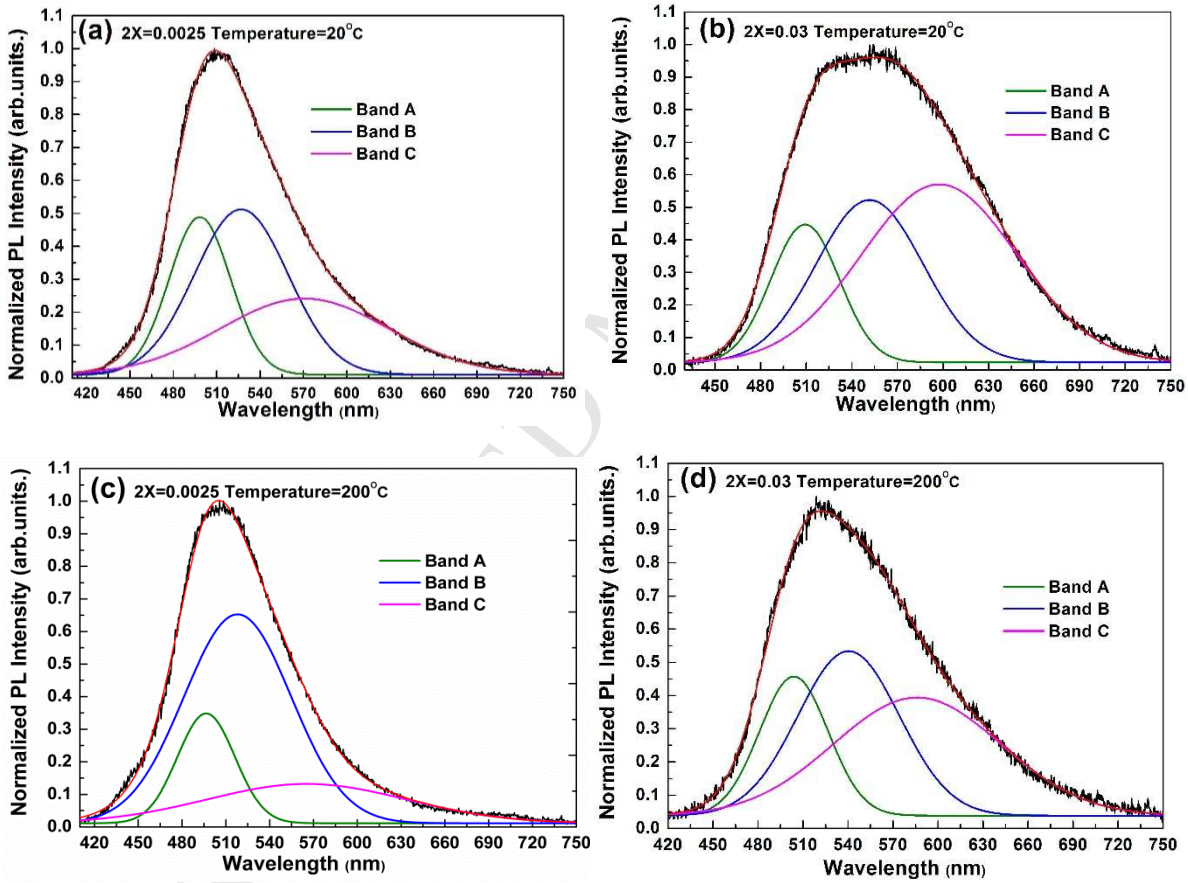

(a) 2X=0.0025
20°C
120°C
150°C
200°C
Intensity (arb.units.)
Wavelength (nm)
(b) 2X=0.03
20°C
120°C
150°C
200°C
Intensity (arb.units.)
Wavelength (nm)
(c) 2X=0.0025
20°C
120°C
150°C
200°C
Normalized PL Intensity (arb.units.)
Wavelength (nm)
(d) 2X=0.03
20°C
120°C
150°C
200°C
Normalized PL Intensity (arb.units.)
Wavelength (nm)

Figure 4.


(a) 2X=0.0025 Temperature=20°C
(b) 2X=0.03 Temperature=20°C
(c) 2X=0.0025 Temperature=200°C
(d) 2X=0.03 Temperature=200°C
Band A
Band B
Band C
Normalized PL Intensity (arb.units.)
Wavelength (nm)

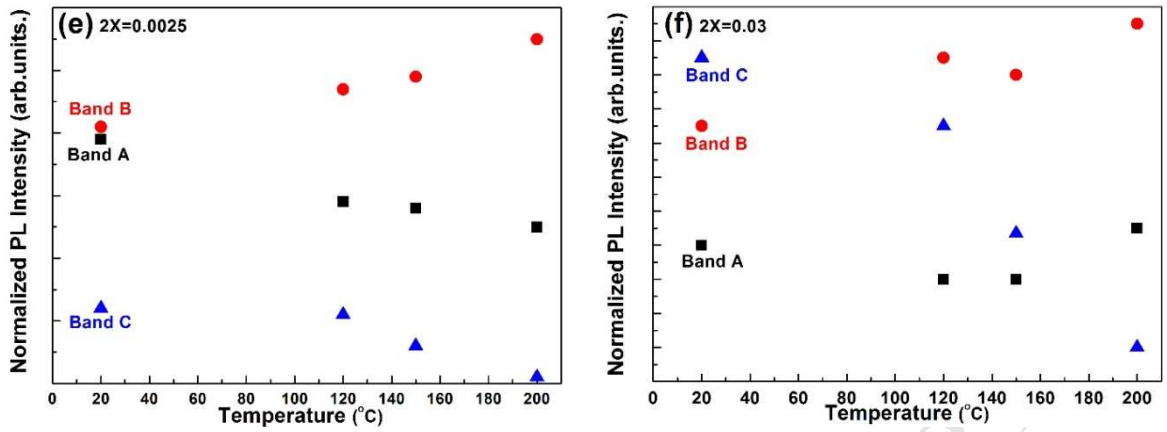
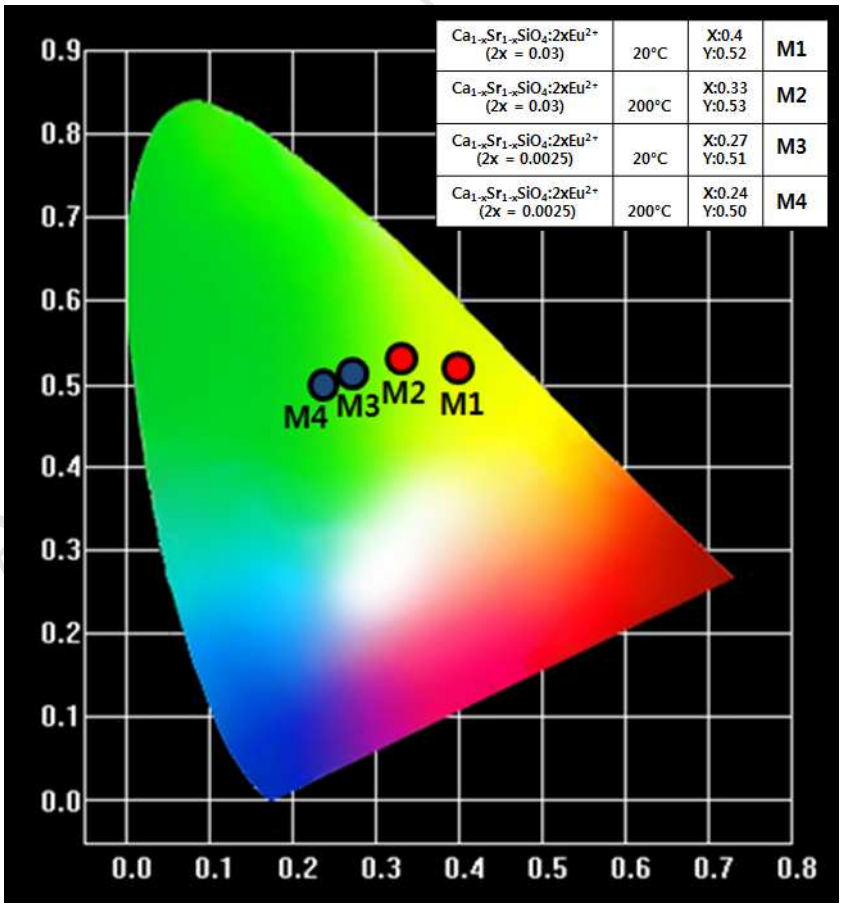

(e) 2X=0.0025
Band B
Band A
Band C
Normalized PL Intensity (arb.units.)
Temperature (°C)
(f) 2X=0.03
Band C
Band B
Band A
Normalized PL Intensity (arb.units.)
Temperature (°C)

Figure 5.


Ca1-xSr1-xSiO4:2xEu2+ (2x = 0.03) 20°C X:0.4 Y:0.52 M1
Ca1-xSr1-xSiO4:2xEu2+ (2x = 0.03) 200°C X:0.33 Y:0.53 M2
Ca1-xSr1-xSiO4:2xEu2+ (2x = 0.0025) 20°C X:0.27 Y:0.51 M3
Ca1-xSr1-xSiO4:2xEu2+ (2x = 0.0025) 200°C X:0.24 Y:0.50 M4
M4
M3
M2
M1

ACCEPTED MANUSCRIPT

Highlights

- $Ca_{1-X}Sr_{1-X}SiO_4:2xEu^{2+}$ phosphors prepared by the Sol-gel method using PGMS
- Temperature dependent luminescence & structural properties of synthesized phosphor
- Effect of temperature on the variation of color of synthesized phosphor
- This reports the site specific temperature dependent luminescence properties

ACCEPTED MANUSCRIPT

Empir Software Eng
DOI 10.1007/s10664-016-9488-7

An empirical study for software change prediction using imbalanced data

Ruchika Malhotra[1] · **Megha Khanna**[1,2]

© Springer Science+Business Media New York 2017

Abstract Software change prediction is crucial in order to efficiently plan resource allocation during testing and maintenance phases of a software. Moreover, correct identification of change-prone classes in the early phases of software development life cycle helps in developing cost-effective, good quality and maintainable software. An effective software change prediction model should equally recognize change-prone and not change-prone classes with high accuracy. However, this is not the case as software practitioners often have to deal with imbalanced data sets where instances of one type of class is much higher than the other type. In such a scenario, the minority classes are not predicted with much accuracy leading to strategic losses. This study evaluates a number of techniques for handling imbalanced data sets using various data sampling methods and MetaCost learners on six open-source data sets. The results of the study advocate the use of resample with replacement sampling method for effective imbalanced learning.

Keywords Change proneness · Data sampling · Empirical validation · Imbalanced learning · MetaCost learners · Object-oriented metrics

1 Introduction

Software change prediction involves determination of classes which are prone to change in the future releases of the software product by taking into account various design

Communicated by: Jeffrey Carver

✉ Ruchika Malhotra
ruchikamalhotra2004@yahoo.com

Megha Khanna
meghakhanna86@gmail.com

[1] Delhi Technological University, Delhi, India

[2] Delhi, India

Published online: 05 January 2017

metrics. These metrics represent various Object-Oriented (OO) software characteristics such as coupling, cohesion, size etc. Identification of change prone classes is essential as these classes require more attention during maintenance and testing phase of the software as they are probable sources of defects and advancements (Eski and Buzluca 2011; Malhotra and Khanna 2013) i.e. they are sources of defects based changes, advancement bases changes as well as maintenance based changes. Proper allocation of resources to such classes helps in improving the quality of the software product as stringent verification activities such as inspections and reviews can be performed on such classes in the early phases of software development so that defects and probable changes can be identified as early as possible. Such a practice would help in easy management of changes and efficient correction of defects in the initial phases of the software life cycle. However, if these defects and changes are propagated to later phases of the software product life cycle, they become costlier to correct and manage. Furthermore, early determination of change prone classes facilitates preventive design measures, which may be taken by software professionals so that future errors may be avoided in these classes (Koru and Tian 2005; Koru and Liu 2007). Also, activities like refactoring can be directed on change-prone classes so that these classes are better adept at handling future changes and the effect of changes made to these classes may be localized (Elish and Al-Khiaty 2013). Thus, development of models for effective determination of change prone classes is important in order to release and maintain good quality software products at optimum costs.

A number of studies in the past (Bieman et al. 2001; Koru and Tian 2005; Eski and Buzluca 2011; Lu et al. 2012; Elish and Al-Khiaty 2013; Malhotra and Khanna 2013) have validated statistical as well as Machine-Learning (ML) methods for determination of change prone classes on various commercial and open-source data sets. In order to develop a useful change prediction model, a learning technique requires an efficient training data set, which have sufficient number of both change-prone and not change-prone classes so that the model can effectively learn to identify them. However, in real world a number of data sets are imbalanced in nature i.e. a majority of classes belong to a particular category, with very few instances of the other category, which is generally of more interest (change prone in our case). Learning from such imbalanced data sets leads to higher misclassifications for the minority class. This is because of lack of information available about the minority class. Such models are rarely of any practical use as the important change-prone classes are neglected and may not be properly identified. This would lead to improper testing and maintenance plans as sufficient resources would not be allotted to the unidentified change prone classes leading to poor quality software products. Moreover, such products would be costlier to maintain and manage and may not be completed under tight schedule deadlines. This study deals with two specific approaches for efficiently learning from imbalanced data sets while developing change prediction models: data sampling approaches and cost-sensitive classification.

(i) Data Sampling Approach: This approach involves modification of training data in such a manner so as to provide adequate number of training instances for both change-prone and not change-prone classes. For example, a researcher can oversample the number of minority instances to increase the ratio of minority examples and improve the classifier's performance (Chawla et al. 2002; Lopez et al. 2013).

(ii) Cost-Sensitive Classification: This approach involves giving appropriate weights to various misclassification errors so that the model learns in such a fashion that it reduces the cost of misclassifications (Domingos 1999; Lopez et al. 2013).

Apart from these methodologies, this study also adopts stable performance metrics which should be evaluated when we are learning through imbalanced data sets. Thus, the research questions (RQ) explored in the study are as follows:

RQ1a: Does the performance of different ML methods in this study significantly improve by using various sampling approaches for Imbalanced Learning Problem (ILP) in change prediction?
RQ1b: Which sampling approach performs best amongst different sampling techniques analysed in the study?
RQ2: What is the effectiveness of different MetaCost learners (Domingos 1999) for improving learning through imbalanced data?
RQ3: What is comparative performance of the best sampling approach and MetaCost learner for ILP?

In order to evaluate the above RQ's we empirically evaluate six widely used open-source data sets using several ML methods. The ILP was addressed by using three data sampling approaches namely, Resampling with replacement, Synthetic Minority Oversampling Technique (SMOTE), and Spread Subsample and by using MetaCost learners for cost-sensitive learning. Furthermore, the study assesses the performance using several stable performance measures like G-mean, Receiver Operating Characteristic (ROC) analysis and balance. The study also compares these stable performance metrics with traditional metrics such as "accuracy", "recall" and "precision". This study develops change prediction models using ten-fold cross validation and inter-release validation. Moreover, the study performs statistical analysis in order to compare various methods for handling ILP and evaluate their effectiveness. The results of the study advocate the use of the above mentioned approaches for developing models using imbalanced data. Furthermore, the study supports the use of resample with replacement sampling method as the most effective approach amongst the evaluated approaches.

The organization of the study is as follows: Section 2 briefly overviews the related literature and Section 3 discusses the ILP in detail. Section 4 discusses the various design aspects of the study and Section 5 describes the experimental design. Section 6 states the research methodology. Section 7 states and analyses the results of the study using ten-fold cross validation and Section 8 discusses the results specific to each RQ using inter-release validation. Section 9 discusses the various threats to validity of the study and Section 10 states the conclusions of the study and provides pointers for future work.

2 Related Work

This section discusses the related work of the study and is organized into three sub-sections. The first section states the studies related to change prediction. The second section reports the various approaches for solving the ILP. The third section states the studies related to ILP on various real-life and software engineering applications especially software defect prediction and change prediction.

2.1 Change Prediction Studies

There are multiple reasons for a change in a software which include existence of defects, change in requirements or upgradation of software. Various researchers in the past have developed models for predicting change prone nature of a class. A study by Bieman et al. (2001) validated an OO system which was commercial in nature to evaluate the relationship between design structure and software changes. In order to determine the change prone classes, the study used design patterns, size metrics and inheritance metrics. Studies by Koru and Tian (2005) and Koru and Liu (2007) use various OO metrics which are indicators of inheritance, size, coupling and complexity attributes of a software on two widely used open-source data sets (KOffice and Mozilla) in order to ascertain change-prone classes. A study by Zhou et al. (2009) validated the confounding effect of size of a class on the relationship between OO metrics and change-prone nature of a class. The study used two releases of Eclipse software, which is a framework of tools for building and deploying software projects. Lu et al. (2012), Eski and Buzluca (2011) and Malhotra and Khanna (2013) evaluated a number of open-source Java projects for ascertaining the nature of change prone classes using various OO metrics which represent size, cohesion, coupling and inheritance attributes of a software. A recent study by Elish and Al-Khiaty (2013) successfully used evolution-based metrics along with OO metrics for determining the change prone nature of a class on two open-source software data sets. A study by Romano and Pinzger (2011) predicted change-prone Java interfaces on Hibernate and Eclipse open-source software data sets using cohesion metrics and other OO metrics. Another study which successfully evaluated change prone classes using OO metrics on two large open source software projects was conducted by Giger et al. (2012). These studies confirm the effective use of change prediction models in various commercial and open source software data sets. Software practitioners can use these models for efficient resource allocation of limited project resources such as its budget, effort and time to the identified change prone classes so that good quality software products are delivered. However, use of imbalanced data sets may pose problems in developing effective change prediction models and the next study discusses various approaches for handling the ILP.

2.2 Approaches for Handling ILP

It is important for ILP to be effectively handled in order to provide efficient learning for model development. A number of issues are associated with ILP. Visa and Ralescu (2005) briefly overview the field of imbalanced learning to discuss the nature and issues associated with imbalanced data. They attribute the ineffective performance of various learning methods to three primary reasons i.e. accuracy, class distribution and error costs. Studies by He and Garcia (2009), Weiss (2004), Zhang and Li (2011) and Lopez et al. (2013) discuss various state of the art methods for developing effective models using imbalanced data. A broad categorization of the various methods suggested by these studies is as follows: (i) use of sampling methods, (ii) use of cost sensitive methods, (iii) use of kernel-based and active learning methods, (iv) use of ensemble learners, (v) use of appropriate evaluation metrics (Weng and Poon 2008; Jeni et al. 2013; Bekkar et al. 2013), (vi) use of knowledge/human interaction, (vii) data segmentation, (viii) use of non-greedy techniques for searching, (ix) use of an effective inductive bias and (x) use of additional methods such as one-class learning methods or novelty detection methods amongst other additional methods. It may be noted that there are certain classification methods that do not assume that the data is balanced is nature. One such method is Mahalanobis-

Taguchi method (Hirohisa et al. 2006; Su and Hsiao 2007), which has been used for developing classification models on imbalanced data sets in quality engineering field. However, these methods have been rarely use in the software engineering field. A prime reason for such a trend is the difficulty in determination of a suitable threshold to conduct efficient classification process while learning from one class samples (Su and Hsiao 2007).

2.3 ILP on Various Real-Life and Software Engineering Problems

A number of various real life applications face ILP, which include sentiment analysis (Xu et al. 2015), fraud detection (Phua et al. 2004), video mining (Apandi et al. 2011), text mining (Munkhdalai et al. 2015), churn prediction (Bekkar et al. 2013) and various bioinformatics applications (Yang et al. 2014) amongst many others. However, in terms of software engineering problems, defect prediction using imbalanced learning data is widely explored by a number of researchers. A study by Wang and Yao (2013) investigated threshold mining, resampling techniques and ensembles techniques for software defect prediction using imbalanced data sets. A number of other studies including Shatnawi (2012), Kamei et al. (2007), Seiffert et al. (2014), Liu et al. (2006) have explored the use of different sampling methods in order to improve the performance of software defect prediction models.

Seliya and Khoshgoftaar (2011) analysed different cost-sensitive learning techniques using decision trees for developing defect prediction models on software data sets. They also took into account misclassification cost as an important parameter for model training. Studies by Weiss et al. (2007), Galar et al. (2012) and Rodriguez et al. (2014) analysed cost-sensitive, sampling and ensemble learning approaches and compared them for developing effective software defect prediction models using imbalanced data. A study by Gao et al. (2015) suggest the use of feature selection along with sampling approaches to learn effectively from imbalanced defect data sets.

Though a number of studies assessed and improved the development of software defect prediction models, only one study by Tan et al. (2015) evaluated change classification models, which is a related area of defect prediction. This study assessed the comparative performance of resampling techniques and updateable classification method for imbalanced learning and ascertained that the performance of developed models improve with the use of these techniques.

This study uses two methods (use of sampling approaches and cost-sensitive methods) for handling ILP. To the best of author's knowledge, this study is a pioneer in statistically evaluating the performance of sampling methods and MetaCost learners on imbalanced data sets for developing change prediction models. Moreover, the study uses effective performance evaluators as discussed in literature for evaluating models created using imbalanced data. Also, the models are compared using certain traditional metrics, which helps in comparing and assessing the imbalanced learning problem from various aspects.

3 The Imbalanced Learning Problem

Imbalanced Learning problem is well recognized in the literature where a data set has large mismatch between the number of instances of different categories of classes. This study deals with binary class imbalance problem where a data set has only two categories of classes namely change prone and not change prone. We generally find that change-prone classes are

under-represented in data sets as software data sets follow Pareto principle. This means that bulk of changes and defects in a data set originate from only 20% of classes or modules (Koru and Tian 2005). Thus, it is important to develop software quality models to identify change prone classes as these classes should be efficiently designed and rigorously tested to improve the quality of the software.

An effective classification model should be able to detect both change-prone and not change-prone classes with high accuracies. However, due to imbalanced nature of the training data, it is often the case when not change prone classes are recognized with higher accuracies but change prone classes are not that well recognized giving lower accuracy rates. However, since the datasets are highly imbalanced, this lower recognition of change prone classes is ignored and the classifiers exhibit good overall accuracy rates. Such classifiers are imbalanced but are still termed as good classification models. However, this scenario is quite unfavourable and may lead to erroneous judgements leading to high losses and bad reputation for the software organization.

Let us understand the scenario with the help of an example. Suppose we have a data sets with 500 data points, where each data point is a collection of OO metrics and the change prone or not change prone nature of a class. The goal is to develop an accurate and unbiased model which detects both change prone and not change prone classes correctly. However, the data set is imbalanced with only 10% of change prone classes and the other not change prone in nature. A model trained on such a data set would show correct and accurate predictions for not change prone classes (close to 100%) but very low accurate predictions for change prone classes ranging from 0 to 10% (He and Garcia 2009). Thus, assuming 10% accuracy for change prone classes only 5 out of 50 change prone classes are accurately predicted. Such low prediction accuracy for change prone classes is not advisable as it would lead to a bad quality software product. This is because change prone classes require allocation of sufficient resources so that they can be effectively designed, scrutinized, verified and tested. These classes should be properly handled in order to prevent occurrences of defects and changes in future versions of the software. A low accuracy for such classes would mean negligence of these classes which would adversely affect the quality of the software. Similarly, if we had assumed that the prediction accuracy for not change prone classes is low, it would mean that a number of not change prone classes are allocated ample resources but these resources are wasted as they are not required for managing such classes. This would lead to overshooting of schedule deadlines and budget overruns leading to poor reputation for the software organization. Thus, it is crucial for software practitioners to effectively ILP where both types of misclassifications are given equal importance so that good quality products are delivered within the allocated time and budget.

4 Experimental Design

This section states and explains the various design decisions of the study.

4.1 Variables Selection

In order to develop change prediction model, the study uses 18 metrics which represent various OO characteristics. These metrics are used to predict a software

quality attribute i.e. change proneness. The sections below briefly describe the independent and dependent variables of the study.

4.1.1 Independent Variables: OO Metrics

OO metrics are important to characterize various software properties like its cohesiveness, size, cohesion, reusability etc. These metrics are used to understand the product and manage it. The various OO metrics used in the study are briefly described below.

- Chidamber and Kemerer (CK) metrics suite: This suite comprises of six popularly used metrics namely Response For a Class (RFC), Depth of Inheritance Tree (DIT), Number of Children (NOC), Weighted Methods of a Class (WMC), Lack of Cohesion in Methods (LCOM) and Coupling Between Objects (CBO) (Chidamber and Kemerer 1994). It is a widely used metric suite used by a number of studies for developing effective software quality prediction models (Elish and Al-Khiaty 2013; Giger et al. 2012; Singh et al. 2009).
- Quality Model for Object-Oriented Design (QMOOD) metrics suite: Certain metrics from this metrics suite namely Number of Public Methods (NPM), Method of Functional Abstraction (MFA), Measure of Aggression (MOA), Cohesion Among Methods of a class (CAM) and Data Access Metric (DAM) are used in the study for developing an efficient software change prediction model. There have been certain previous studies in literature which have already used this metrics suite for quality prediction tasks (Eski and Buzluca (2011); Olague et al. (2007)).
- Other Metrics: The study also evaluates a set of other miscellaneous metrics which include Afferent Coupling (Ca) and Efferent Coupling (Ce) metrics proposed by Martin (2002). Some other commonly used metrics in literature such as Lines of Code (LOC), Coupling Between Methods of a Class (CBM), Average Method Complexity (AMC), LCOM3 (proposed by Henderson-Sellers 1996) and Inheritance Coupling (IC) are also used in the study for developing change prediction models.

In order to refer to detailed definition of the metrics used in the study, refer to http://gromit.iiar.pwr.wroc.pl/p_inf/ckjm/metric.html.

4.1.2 Dependent Variable

This study analyzes a binary dependent variable i.e. change proneness, which symbolizes the change prone nature of a class. In order to comprehend the change prone nature of a class, we analyze two different versions of the same software, say an old version and a more recent new version. A class is termed as change prone if certain LOC have been added, deleted or modified in the corresponding class, in the new version of the software as compared to the old version, otherwise it is termed as not change prone. Previous studies have also characterized the change proneness dependent variable in the same way (Lu et al. 2012; Malhotra and Khanna 2013) as taking into account the change in LOC is a practical measure for determining the binary variable change proneness. The objective of the study is to build a prediction model that predicts the outcome label for unforeseen classes.

4.2 Subject Selection

The study uses three application packages of the Android data set and three common software developed by Apache. Android is an extremely popular operating system for mobile devices, which is open source in nature and is developed by Google. This study uses three application packages of Android system namely Calendar, Bluetooth and MMS which are used for application specific functions. The data collected by Android systems is analyzed from two different versions of Android i.e. Ice Cream Sandwich and Jelly Beans. Furthermore, the study also evaluates three apache data sets namely IO, Net and Log4j. Apache Net implements a number of internet protocols on the client side. Apache IO implements a number of utilities which help in input-output functionality. These utilities may include input, output, comparator, file monitor etc. Apache Log4j software is used to embed logging statements at runtime for debugging purposes. Both Android and Apache uses GIT as the version control system. Moreover, the open source nature of the selected data sets aids the replicability of the study. Wide use of Android and apache data sets enhances the external validity of the study as the results of the study can be applied effectively in similar scenarios.

4.3 Statistical Test Selection

In order to assess and validate various hypothesis of the study, we use Friedman statistical test which is followed by Wilcoxon post-hoc test. These tests are non-parametric and thus can be effectively used without much inconvenience as they do not depend on a number of assumptions of underlying data such as data normality, homogeneity of variances etc., which are mandatory for parametric tests (Demšar 2006). Moreover, as reported by Lessmann et al. (2008), very few studies evaluate the results statistically while comparing various developed models. Thus, statistical analysis is important to strengthen the conclusions of the study.

The Friedman test is used to ascertain whether there is any significant differences in the change prediction models developed using various sampling methods (RQ1). Furthermore, the Friedman test assigns a mean rank to different sampling methods depending on their performance metrics. If the results of Friedman test are significant, we should perform post-hoc Wilcoxon test to ascertain pairwise differences amongst different sampling methods. The Wilcoxon test calculates the differences amongst the performance of different pairs of sampling methods for comparison. These computed differences are then allocated ranks in accordance with their absolute values. This study also uses Wilcoxon test for evaluating the best sampling method with MetaCost learners (RQ3).

4.4 Data Collection

This study uses six data sets which are collected using Defect Collection and Reporting System (DCRS) tool. The tool was developed in Java language by undergraduate students of Delhi Technological University. The tool is used to collect data from various open source repositories with the precondition that the repository uses GIT as its versioning control system (Malhotra et al. 2014). In order to analyze the source code of a particular data, the DCRS tool extracts change-logs between two consecutive versions of the software data set. These change logs consist of change records which specify change information including change description, its identifier, file listing of the modified files, commit timestamp, listing of changed lines of code etc.

Each data set consists of a number of data points, where each data point corresponds to OO metrics of a class and its change statistics. The change statistics include the number of insertions, deletions, modifications and the total change in a specific class of a software when two versions of a software are compared. It also has a binary variable which depicts the change prone nature of a class. This variable named ALTER is "yes" if there is a change in certain LOC of a class and "no" otherwise. In order to compute the values of different OO metrics, the DCRS tool uses another open source tool CKJM (http://gromit.iiar.pwr.wroc.pl/p_inf/ckjm/metric.html).

This study uses three application packages of Android OS and three Apache software, which use Git as their versioning control system. In order to analyze change, we use only Java class files, ignoring all other files related to media, layout etc. Table 1 lists the details of each software data set which includes the two versions analyzed, number of data points in each data set and the percentage of change prone classes. The data points correspond to the number of common classes in the two analyzed versions of each data set. A number of studies in literature have analyzed imbalanced data of varying nature i.e. data sets which are majorly imbalanced, moderately imbalanced to data sets which are mildly imbalanced (Hulse et al. 2009; Lopez et al. 2013; Liu et al. 2006). This study uses data sets in which percentage of change prone classes is in the range of 6–37%, indicating the use of highly imbalanced to moderately imbalanced data sets for developing change prediction models.

4.5 Performance Measures Selection

In order to evaluate models developed using imbalanced data, we need appropriate performance metrics. Previous studies in literature have criticized the use of traditional performance evaluators like accuracy and precision when evaluating models using imbalanced data sets (He and Garcia 2009; Gao et al. 2015; Menzies et al. 2007b). However, a number of studies in literature have discussed the favorability of using other robust performance metrics for assessing models developed using imbalanced data. Studies by Harman et al. (2014), Kubat and Matwin (1997) and He and Garcia (2009) support the use of G-mean as a performance metric for imbalanced data sets. Certain other studies by Menzies et al. (2007a) and Li et al. (2012) advocate balance performance metric as an efficient evaluator. Moreover, studies by Lessmann et al. (2008), Shatnawi (2012) and He and Garcia (2009) propose the use of ROC analysis, where Area Under ROC Curve (AUC) can be used as an effective indicator for evaluating models developed using imbalanced data. AUC is robust in handling skewness in class distributions and the unequal cost of misclassification errors (Fawcett 2006). Apart from

Table 1 Software data set details

Software name	Versions analyzed	No. of data points	% of changed classes
Android calendar	4.3.1–4.4.2	106	19%
Android bluetooth	4.3.1–4.4.2	72	19%
Android MMS	2.3.7–4.0.2	195	30%
Apache IO	2.3–2.4	198	6%
Apache net	3.0–3.1	240	37%
Apache Log4j	1.2.16–1.2.17	350	25%

the above mentioned performance metrics, the study also evaluates traditional performance metrics namely accuracy, recall and precision.

The various performance metrics used in the study can be evaluated using confusion matrix which is stated as follows. A confusion matrix for a problem with two classes namely change prone and not change prone is shown in Table 2. The positives are considered as change prone classes and the negatives are not change prone classes.

Table 3 describes the various performance metrics assessed in the study along with the formula for computing them.

5 Experimental Design

This section describes the experimental design used in the study. Figure 1 presents the pictorial representation of the experimental design. The experiment is conducted in three phases which are described as follows.

5.1 Data Pre-processing and Feature Selection

This phase involves three steps. The first step focuses on analysing the descriptive statistics of different OO metrics for each data set. This helps in assessing data characteristics of each data set. The next step focuses on identification of outliers from each data set. Outlier identification is critical in order to develop effective models which are not influenced by extreme variability in data points. This study removed outliers using the Inter Quartile range filter, which computes outliers by assessing the difference between upper quartile (Q_3) and lower quartile (Q_1). A specific data point is termed as an outlier if any of the OO metric i.e. independent variables (mentioned in section 4.1.1) in the data point is considered as an outlier. Outliers from each of the six data sets are identified and removed in order to provide effective data for model development.

The final step of this phase involves use of Correlation based Feature Selection (CFS) method in order to eliminate noisy and redundant independent variables from each data set. According to Hall (2000), the CFS method helps in identifying those independent variables which are highly correlated with the change in the class but not amongst each other. The set of independent variables extracted after application of the CFS method are important as they help in reducing the dimensionality of data sets. The CFS method evaluates an exhaustive possible combination of OO metrics, to select the best subset of OO metrics for change prediction in each corresponding data set. It may be noted that there is no single best technique for selecting features in a data set (Gao et al. 2015; Hall and Holmes 2003). A study by Hall and Holmes (2003), evaluated six attribute selection methods on 15 data sets and confirmed that though there is no best method for attribute selection, the CFS method shows good results. Furthermore, an extensive review conducted by Malhotra (2015) of 64 defect prediction

Table 2 Confusion matrix

	Predicted positive	Predicted negative
Actual positive (Change Prone)	True Positive (TP)	False Negative (FN)
Actual negative (Not Change Prone)	False Positive (FP)	True Negative (TN)

Table 3 Performance metrics

Performance metric	Description	Formula
Accuracy	It is defined as the percentage of classes which are correctly predicted. These classes include both positive and negative classes.	$\frac{TP+TN}{TP+FP+FN+TN}$ *100
Recall	It is defined as the percentage of correctly predicted positives.	$\frac{TP}{TP+FN}$ *100
Specificity	It is defined as the percentage of correctly predicted negatives.	$\frac{TN}{FP+TN}$ *100
Precision	It is defined as the percentage of correctly predicted positives amongst total number of predicted positives.	$\frac{TP}{TP+FP}$ *100
G-mean	It is the geometric mean of both positive as well as negative accuracy.	$\sqrt{\frac{TP}{TP+FP} * \frac{TN}{TN+FN}}$
Balance	It depicts the Euclidean distance between a pair of (Recall, PF), where PF corresponds to Probability of False Alarm to that of an optimal value of Recall =1 and PF = 0. PF corresponds to the percentage of incorrectly predicted change prone classes amongst actual not change prone classes.	$1-\sqrt{\frac{(0-(PF/100))^2+(1-(PD/100))^2}{2}}$ *where* $PF = \frac{FP}{FP+TN}$ *100
Area under ROC curve (AUC)	ROC represents a plot between 1-specificity values on the x-axis and recall values on the y-axis. A model is considered favorable if it has higher area under the ROC curve.	

studies from 1991 to 2013 points out that CFS was the most commonly used feature selection in those studies. Since a large number of studies in literature have successfully applied the CFS method for feature selection (Arisholm et al. 2010; Carvalho et al. 2010; Malhotra and Khanna 2013) while developing predictive models using ML techniques, we use this method.

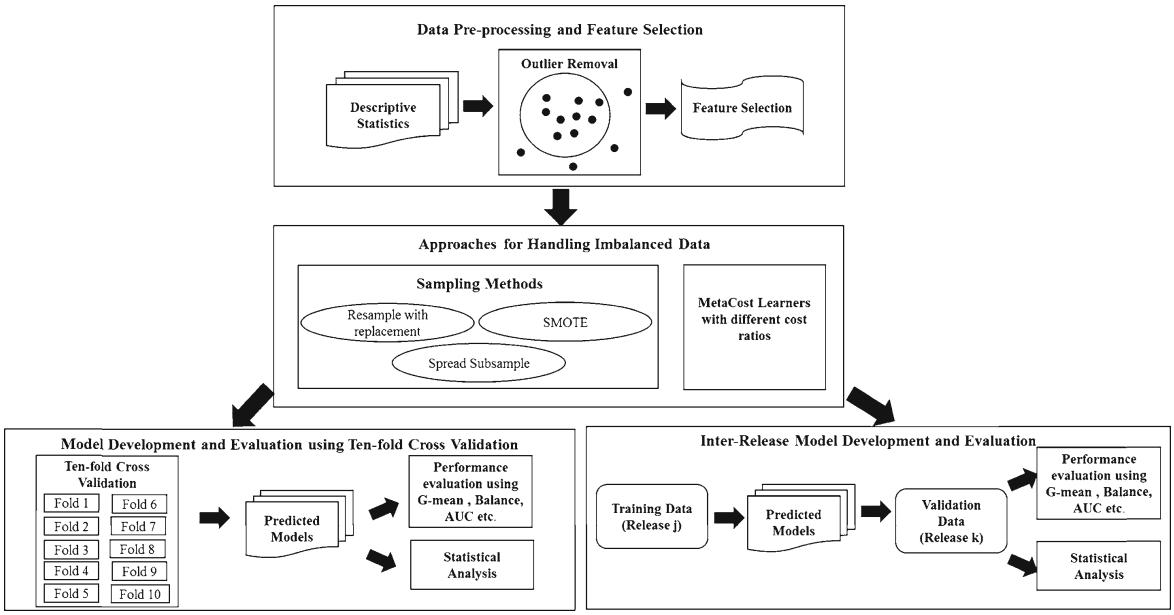


Fig. 1 Experimental design

5.2 Approaches for Handling Imbalanced Data

This phase includes application of various methods in order to treat imbalanced data and provide efficient and balanced data for training using sampling methods or provide cost-sensitization to various misclassification errors so that the model minimizes the various errors of misclassification. Thus, this phase involves use of sampling methods or MetaCost learners for efficiently handling imbalanced data and developing effective software quality models. The study evaluates three sampling methods and MetaCost learners.

5.3 Model Development and Evaluation

This phase involves development of change prediction models using six ML techniques on each of the six data sets of the study. In order to develop change prediction models, we use ten-fold cross validation technique (Stone 1974). This technique functions by performing division of all the data points of a specific data set into ten subsets. The technique then performs ten iterations by using nine subsets for training and the tenth subset for validation until we use each of the ten subsets once for validation purposes. We use ten-fold cross validation technique as it reduces validation bias (Carvalho et al. 2010; Pai and Dugan 2007). Apart from using ten-fold cross validation technique, the study also uses inter-release model validation for developing change prediction models. For developing models using inter-release validation, we train the model using a specific version of the data set and validate the model on another version of the data set. It should be noted that outliers are only removed from the training data sets while developing inter-release validation models. The next step in this phase is to evaluate the performance of the developed change prediction models. For this purpose, we use three effective performance evaluators i.e. G-mean, balance and AUC as discussed in Section 4.5. Furthermore, these performance metrics are compared with three traditional performance metrics i.e. accuracy, recall and precision. In order to increase the conclusion validity of the study, we also perform statistical testing to compare the performance of different approaches for handling imbalanced data.

6 Research Methodology

The study uses three different sampling methods along with MetaCost learners for handling ILP. The change prediction models on each data set are developed using six ML techniques. The ML techniques used in the study include Multilayer Perceptron (MLP), Random Forests (RF), Naïve Bayes (NB), Adaboost (AB), Logitboost (LB) and Bagging (BG). The study uses the default parameter settings of WEKA tool (http://www.cs.waikato.ac.nz/ml/weka/) for simulation.

6.1 Machine Learning Techniques

This section states a brief description of the various ML techniques used in the study along with the parameters used for each ML technique.

MLP are artificial neural networks which map a set of inputs from the input layer to the desired set of outputs using a number of hidden layers (Haykin 2004). They are feedforward networks which simulate the biological neurons. Backpropagation is the most common mechanism to train MLP where two passes take place i.e. forward and backward through the network.

In the forward pass, the inputs are applied and the output is produced, which is the actual response of the network. During the process, the synaptic weights of the network are adjusted. The backward pass propagates the error signal through the network which is the difference of actual output and the desired output. The weights of the network are again readjusted so that the actual response becomes closer to the desired response. The default parameter settings for MLP used in WEKA are learning rate of 0.3, momentum of 0.2 and a validation threshold of 20.

RF, proposed by Brieman (2001) is a collection of many decision trees. It is an ensemble classifier which outputs the class which is the mode of the output of all individual decision trees. This property helps in overcoming the overfitting characteristic of decision trees. The default parameter settings for RF include 100 decision trees in the random forest.

NB is a probabilistic classifier. It is based on the assumption that various input features are independent and uses Bayes theorem. The classification models developed using NB techniques are easily interpretable and are highly useful when there are large number of input features (Murphy 2006). The default settings of WEKA do not use kernel estimator or supervised discretization.

AB is a boosting method which improves the efficiency of poor classifiers and transforms them into strong classifiers. The algorithm provides a weighted sum of various weak learners to construct an output which is of an efficient boosted classifier. The initial step involves providing equal weights to all the classifiers (Witten et al. 2011). However, with each iteration, the weights are adjusted in such a manner that the algorithm focuses on the hard to learn examples to improve the efficiency of the boosted classifier. The default parameter settings for WEKA tool includes Decision Stump classifier, 10 iterations and a weight threshold of 100.

LB is another boosting method which improves the efficiency of weak learners to construct an efficient classifier. The process involves reweighting of training instances iteratively to produce a strong classifier. The LB technique uses AB technique for additive model and applies the cost functional of the logistic regression technique (Friedman et al. 2000). The default parameter settings for LB technique in WEKA include the Decision Stump classifier, 10 iterations, weight threshold of 100 and a likelihood threshold of −1.79.

BG also known as bootstrap aggregating is a technique which constantly improves the developed classification models by creating a number of versions of training set. Breiman (1996) proposed creating bootstrap duplicates of the training set, where sampling is done with replacement and a new function is trained for each of the training set. In order to predict a class, the result of majority is output. The default settings for BG technique in WEKA include REP tree as the classifier, a bag size percent of 100 and 10 iterations.

6.2 Data Sampling Methods

This section briefly describes the three sampling methods used in the study.

6.2.1 Resample with Replacement Method

The resample method is based on the bootstrapping approach. This method creates a number of subsamples, which are random in nature by using with replacement method. This is done in order to influence the original class distribution and achieve a more uniform class distribution. In with replacement method, the minority class is oversampled in order to achieve a uniform number of classes of both change prone and not change prone nature. Since, each time a different sample is

created in both the methods which is random in nature, we perform ten iterations for these methods and report the average results. The method biases the data set to produce uniform ratio of both change prone and not change prone instances.

6.2.2 Spread Subsample

This method is used to adjust the spread of the two classes by removing instances of the class which is in majority. This would help in balancing the corresponding ratios of both the classes (Tan et al. 2015). This study evaluates a number of ratios and selects the best ratio which may vary from 1:1–10:1.

6.2.3 SMOTE

SMOTE is an oversampling method where the minority classes are oversampled to increase the number of instances which belong to the minority class. The SMOTE (Synthetic Minority Oversampling Technique) method is used to create artificial instances by oversampling (Chawla et al. 2002). The method selects k-nearest neighbours randomly for oversampling the minority class. This study uses the value of $k = 5$. The next step involves selection of neighbours which is dependent on the amount of oversampling. For example, 200% oversampling would mean, two out of five neighbours are selected. Then, a synthetic sample is forged in the direction of each chosen neighbour. The creation of synthetic sample involves two steps. The first step computes the difference between chosen sample and the selected nearest neighbour. The second step multiplies the computed difference with a random number between 0 and 1, which is then added to the original sample. Such generated samples generalize the decision region of the minority class by creation of new artificial instances (Chawla et al. 2002). This study analyses five different percentage oversampling of the minority classes i.e. we evaluate SMOTE 100, SMOTE 200, SMOTE 300, SMOTE 400 and SMOTE 500.

6.3 MetaCost Learners

As discussed in section 3, both types of misclassification errors i.e. the error which occurs because of incorrect prediction of change prone class as non-change prone or the error which occurs because of misclassifying a non-change prone class as change prone should be minimized and balanced. A MetaCost learner is used to cost-sensitize a specific classification technique by wrapping it up in another procedure which is used to minimize cost. The cost of misclassifying a class ‘a’ as belonging to class ‘b’ is denoted by C(a,b). The MetaCost learners use Bayes optimal prediction in order to minimize the conditional risk (R(a|x)). The risk ascertains the expected cost of classifying class x as belonging to category ‘a’ as shown in Eq. (1). The P(a|x) in Eq. (1) is the probability that a specific class x belongs to class category ‘a’. It is certain that Bayes optimal prediction would lead to lowest possible cost for all the instances x. The space of all examples ‘X’ is partitioned into ‘n’ regions. The regions are such that class ‘a’ is the most optimal prediction in terms of cost in region ‘a’.

$$R\left(a\middle|x\right) = \sum_{n} P\left(a\middle|x\right) C(a,b) \tag{1}$$

The MetaCost learners are such that they modify the labels of training examples so that they are representatives of their optimal class. This is done by using an ensemble of ‘k’ classifiers

for learning and using the vote of each classifier. It is basically a variation of bagging technique with replacement. The labels are allocated to each example on the basis of votes received by it or by taking into account the probability estimates of 'k' models ($k = 10$). Therefore, the allocation of a new class label is dependent on the cost ratio and the probability estimates. The new training set with relabelled examples is then used by the base classification technique to develop a prediction model which is cost sensitive (Seliya and Khoshgoftaar 2011; Domingos 1999).

7 Analysis and Results

This section states the analysis steps performed to evaluate the change prediction models. It describes the results of the study.

7.1 Descriptive Statistics

After analysing the descriptive statistics of all the 18 OO metrics for each of the six data sets, which include the minimum, maximum, mean, standard deviation, 25% percentile and 75% percentile, we summarize the following observations:

- The mean values of NOC (Calendar-0.95, Bluetooth-1, MMS-0.85, Apache IO-0.28, Net-0 and Log4j-0.20) and DIT (Calendar-0.37, Bluetooth-0, MMS-0.19, Apache IO-0.73, Net- 1 and Log4j-0.95) suggest that reusability through inheritance was not much used in the data sets. Cartwright and Shepperd (2000) point out a similar trend in their study.
- Each data set had varying class size in terms of LOC. The class size for Calendar data set varied between 8 and 1507 LOC, for Bluetooth data set 9–718 LOC and 0–882 for MMS data set. For Apache data sets the class sizes were in the range 8–768 LOC for IO data set, 0–2372 for Net data set and 1–1864 for Log4j data set.
- An analysis of LCOM values indicated that the data sets exhibit significantly high LCOM values with maximum LCOM values as 5995 for Calendar, 6903 for Bluetooth, 8128 for MMS, 8001 for Apache IO, 6252 for Net and 7503 for Log4j. The mean values of CAM metric varied in the range 0.37–0.51. The values of MFA metric were very low indicating hardly any use of functional abstraction.

7.2 Outlier Analysis

In order to develop effective change prediction models, it is important to identify and remove outliers in the data. Thus, to provide impartial results, we find all outliers using Inter Quartile range filter of WEKA tool and remove them from each data set. The number of outliers determined in each data set are shown in Table 4.

7.3 Correlation Based Feature Selection Results

The independent variables selected by each data set, after the application of CFS method are shown in Table 5 below. According to the table, the most commonly selected OO metric for predicting change prone classes was LOC. Other commonly used metrics were AMC and

Table 4 Outlier details

Data set	Number of outliers
Android calendar	6
Android bluetooth	7
Android MMS	22
Apache IO	11
Apache net	39
Apache Log4j	65

LCOM3, which were selected in three out of six data sets of the study. RFC, WMC, MOA and CAM were selected by two data sets each, indicating their ability to determine change prone nature of a class.

7.4 Discussion and Evaluation of Ten-Fold Cross Validation Results

RQ1a: Does the performance of different ML methods in this study significantly improve by using various sampling approaches for ILP in change prediction?
RQ1b: Which sampling approach performs best amongst different sampling techniques analysed in the study?

In order to assess the change prediction models developed using various sampling methods we use accuracy, recall, precision, G-mean, balance and AUC performance metrics.

Tables 6, 7, 8, 9, 10, and 11 report the values of different performance metrics as shown by the six ML techniques on each of the data set for each sampling method and the scenario where no sampling method is used. It should be noted that SMOTE and spread subsample values are those which are best for a particular data set. Since, a number of combinations of SMOTE and spread subsample are explored, we choose and compare those which give the best results for a corresponding data set.

According to Table 6, the accuracy values of the scenario where no sampling method is used is better than the one in which we use certain sampling methods for handling imbalanced learning in two data sets namely Calendar and IO, and for few cases in Log4j data set. However, accuracy is not an effective indicator of performance evaluation, specifically in the case of imbalanced data sets. This is discussed in a number of literature studies (He and Garcia 2009; Gao et al. 2015; Weng and Poon 2008). However, in all other data sets, it can be

Table 5 CFS results

Data set	Independent variables selected
Android Calendar	CBO, Ce
Android Bluetooth	WMC, RFC, LOC, CAM
Android MMS	LCOM3, LOC, DAM, MOA, CAM, AMC
Apache IO	NPM, LCOM3, AMC
Apache Net	WMC, NOC, LCOM3, LOC
Apache Log4j	RFC, LOC, MOA, AMC

Table 6 Accuracy results using different sampling methods

ML technique	Calendar				ML technique	Bluetooth			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	71.03	51.89	82.56	85.00	MLP	87.09	81.65	73.91	83.08
RF	80.17	76.76	79.07	83.00	RF	96.62	82.57	82.61	78.46
NB	54.62	55.14	79.07	81.00	NB	83.85	78.90	86.96	83.08
AB	74.45	56.22	82.56	85.00	AB	95.23	78.90	86.96	75.38
LB	77.86	74.59	82.56	85.00	LB	96.61	86.24	82.61	78.46
BG	78.04	69.73	81.40	83.00	BG	91.08	80.73	82.61	83.08
ML technique	MMS				ML technique	IO			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	88.43	86.03	75.17	75.72	MLP	81.55	84.85	88.89	95.72
RF	95.72	89.95	69.66	71.68	RF	98.18	85.71	86.67	91.44
NB	77.86	78.19	69.66	70.52	NB	74.65	78.79	84.44	93.58
AB	79.59	85.05	73.10	69.94	AB	81.44	86.58	82.22	93.05
LB	84.34	85.05	73.10	73.99	LB	88.93	85.71	82.22	94.12
BG	89.88	87.75	76.55	78.61	BG	94.12	84.85	88.89	94.65
ML technique	Net				ML technique	Log4j			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	88.75	77.78	77.47	76.62	MLP	77.89	63.82	84.23	81.40
RF	94.82	85.19	75.27	74.63	RF	88.98	81.14	79.65	78.60
NB	72.45	66.40	74.18	74.13	NB	67.54	59.21	78.95	79.65
AB	78.06	81.75	76.92	77.11	AB	74.60	71.27	82.46	82.46
LB	87.36	83.60	75.27	78.11	LB	79.96	76.32	82.46	82.81
BG	88.76	84.92	76.92	77.61	BG	85.36	79.17	81.40	81.40

Table 7 Recall results using different sampling methods

ML technique	Calendar				ML technique	Bluetooth			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	56.73	51.96	17.65	17.65	MLP	81.15	89.09	72.73	27.27
RF	77.27	67.65	17.65	5.88	RF	98.28	83.64	81.82	27.27
NB	22.36	30.39	29.41	29.41	NB	66.55	70.91	81.82	36.36
AB	67.46	78.43	17.65	17.65	AB	94.48	80.00	81.82	9.09
LB	78.73	67.65	17.65	17.65	LB	97.59	85.45	81.82	27.27
BG	74.55	68.63	11.76	17.65	BG	87.93	78.18	81.82	9.09
ML technique	MMS				ML technique	IO			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	93.23	93.62	51.06	44.68	MLP	94.73	38.18	54.55	27.27
RF	97.21	95.74	44.68	40.43	RF	99.89	60.00	63.64	27.27
NB	86.51	83.69	72.34	70.21	NB	71.54	29.09	45.45	27.27
AB	90.69	96.10	68.09	51.06	AB	88.79	60.00	45.45	27.27
LB	88.60	95.39	61.70	38.30	LB	94.62	58.18	54.55	27.27
BG	92.21	96.10	55.32	40.43	BG	97.58	49.09	63.64	9.09
ML technique	Net				ML technique	Log4j			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	89.05	88.56	45.76	40.68	MLP	64.41	53.51	22.81	17.54
RF	94.10	90.68	62.71	50.85	RF	87.21	74.12	31.58	26.32
NB	56.42	54.66	38.98	32.20	NB	37.79	30.26	10.53	14.04
AB	80.53	94.07	61.02	61.02	AB	61.91	59.21	22.81	24.56
LB	87.68	91.53	54.24	57.63	LB	71.91	71.49	22.81	26.32
BG	87.89	91.53	55.93	55.93	BG	81.91	75.00	17.54	19.30

Table 8 Precision results using different sampling methods

ML technique	Calendar				ML technique	Bluetooth			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	76.39	56.99	75.00	75.00	MLP	89.56	77.78	72.73	50.00
RF	80.16	87.34	42.86	50.00	RF	94.59	82.14	81.82	33.33
NB	52.70	72.09	45.45	41.67	NB	95.94	84.78	90.00	50.00
AB	76.38	57.55	75.00	75.00	AB	94.83	78.57	90.00	14.29
LB	76.22	83.13	75.00	75.00	LB	95.04	87.04	81.82	33.33
BG	78.21	74.47	66.67	50.00	BG	92.03	82.69	81.82	50.00
ML technique	MMS				ML technique	IO			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	85.37	87.13	64.86	56.76	MLP	74.50	95.45	100	100.00
RF	95.79	90.30	53.85	47.50	RF	96.53	75.00	77.78	27.27
NB	79.54	84.59	52.31	47.14	NB	76.35	61.54	83.33	42.86
AB	81.57	84.42	57.14	45.28	AB	76.97	78.57	71.43	37.50
LB	84.98	84.86	58.00	52.94	LB	84.82	76.19	66.67	50.00
BG	90.13	87.42	66.67	67.86	BG	91.02	79.41	87.50	100.00
ML technique	Net				ML technique	Log4j			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	87.46	78.57	75.00	66.67	MLP	86.11	67.4	100.00	62.50
RF	95.03	86.29	61.67	57.69	RF	89.58	86.22	48.65	44.12
NB	79.76	86.58	67.65	61.29	NB	87.23	71.88	40.00	47.06
AB	75.63	80.14	65.45	61.02	AB	81.50	78.03	68.42	66.67
LB	85.95	83.72	64.00	64.15	LB	84.02	79.13	68.42	68.18
BG	88.46	85.38	67.35	63.46	BG	86.72	81.82	62.50	61.11

Table 9 G-mean results using different sampling methods

ML technique	Calendar				ML technique	Bluetooth			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	0.72	0.52	0.78	0.80	MLP	0.88	0.82	0.74	0.66
RF	0.80	0.78	0.59	0.65	RF	0.97	0.83	0.83	0.53
NB	0.54	0.60	0.62	0.59	NB	0.87	0.79	0.87	0.66
AB	0.75	0.55	0.79	0.80	AB	0.95	0.79	0.87	0.34
LB	0.79	0.75	0.79	0.80	LB	0.96	0.86	0.82	0.53
BG	0.78	0.69	0.74	0.65	BG	0.91	0.81	0.82	0.65
ML technique	MMS				ML technique	IO			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	0.89	0.85	0.71	0.68	MLP	0.83	0.89	0.93	0.98
RF	0.96	0.90	0.64	0.61	RF	0.98	0.81	0.83	0.51
NB	0.79	0.74	0.66	0.64	NB	0.76	0.71	0.84	0.64
AB	0.81	0.86	0.68	0.61	AB	0.82	0.83	0.78	0.59
LB	0.85	0.85	0.68	0.65	LB	0.89	0.82	0.76	0.69
BG	0.90	0.88	0.73	0.74	BG	0.94	0.83	0.88	0.97
ML technique	Net				ML technique	Log4j			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	0.89	0.77	0.77	0.72	MLP	0.79	0.64	0.91	0.72
RF	0.95	0.85	0.71	0.68	RF	0.89	0.82	0.64	0.61
NB	0.74	0.68	0.72	0.68	NB	0.74	0.63	0.57	0.62
AB	0.79	0.83	0.73	0.72	AB	0.76	0.72	0.76	0.75
LB	0.87	0.84	0.71	0.73	LB	0.81	0.77	0.75	0.76
BG	0.89	0.85	0.74	0.72	BG	0.85	0.79	0.72	0.71

Table 10 Balance results using different sampling methods

ML technique	Calendar				ML technique	Bluetooth			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	66.55	51.88	41.76	41.76	MLP	85.16	80.11	73.84	48.42
RF	79.24	75.59	41.62	33.44	RF	96.37	82.53	82.56	47.98
NB	43.34	49.73	49.71	49.73	NB	76.23	77.48	85.86	54.70
AB	71.32	47.47	41.76	41.76	AB	94.85	78.86	85.86	35.24
LB	75.67	74.20	41.76	41.76	LB	96.49	86.22	82.56	47.98
BG	76.28	69.83	37.60	41.71	BG	89.95	80.59	82.56	35.70
ML technique	MMS				ML technique	IO			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	87.43	77.65	64.15	59.87	MLP	77.62	56.29	67.86	48.57
RF	95.32	83.45	58.78	56.26	RF	97.56	71.37	73.95	48.47
NB	76.23	73.25	70.29	70.42	NB	70.52	49.70	61.37	48.55
AB	76.83	71.80	71.55	61.76	AB	79.49	71.49	61.21	48.53
LB	83.62	72.87	68.97	55.46	LB	86.96	70.16	67.26	48.56
BG	89.42	77.94	67.04	57.57	BG	93.16	63.89	74.20	35.72
ML technique	Net				ML technique	Log4j			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	88.54	70.49	61.30	57.63	MLP	73.45	45.42	62.38	41.67
RF	94.45	81.83	70.50	63.56	RF	88.65	51.26	79.88	47.57
NB	67.67	66.43	56.39	51.69	NB	55.80	36.67	49.98	39.15
AB	77.10	72.29	70.35	70.15	AB	70.82	45.38	68.84	46.61
LB	87.11	78.24	66.03	68.58	LB	77.85	45.38	75.83	47.85
BG	88.47	80.63	67.51	67.43	BG	84.61	41.67	78.75	42.89

Table 11 AUC results using different sampling methods

ML technique	Calendar				ML technique	Bluetooth			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	0.75	0.56	0.51	0.53	MLP	0.87	0.88	0.77	0.79
RF	0.87	0.82	0.55	0.61	RF	0.97	0.88	0.89	0.70
NB	0.59	0.56	0.56	0.58	NB	0.91	0.90	0.95	0.85
AB	0.80	0.61	0.46	0.43	AB	0.96	0.83	0.83	0.66
LB	0.84	0.76	0.47	0.49	LB	0.96	0.91	0.83	0.67
BG	0.84	0.79	0.54	0.54	BG	0.95	0.88	0.89	0.77
ML technique	MMS				ML technique	IO			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	0.90	0.88	0.79	0.80	MLP	0.88	0.85	0.84	0.74
RF	0.98	0.92	0.780	0.61	RF	1.00	0.92	0.81	0.67
NB	0.82	0.82	0.79	0.79	NB	0.88	0.82	0.85	0.72
AB	0.87	0.89	0.78	0.76	AB	0.93	0.87	0.84	0.66
LB	0.91	0.89	0.82	0.79	LB	0.98	0.88	0.86	0.70
BG	0.96	0.90	0.82	0.79	BG	0.99	0.89	0.74	0.67
ML technique	Net				ML technique	Log4j			
	Resample	SMOTE	Subsample	No Sample		Resample	SMOTE	Subsample	No Sample
MLP	0.89	0.85	0.84	0.77	MLP	0.80	0.69	0.67	0.67
RF	0.98	0.92	0.83	0.80	RF	0.96	0.88	0.62	0.61
NB	0.81	0.81	0.78	0.78	NB	0.68	0.62	0.58	0.62
AB	0.84	0.87	0.80	0.78	AB	0.79	0.78	0.64	0.66
LB	0.93	0.87	0.81	0.79	LB	0.87	0.82	0.65	0.68
BG	0.95	0.90	0.81	0.78	BG	0.93	0.85	0.65	0.68

seen that in majority of the cases, the performance of change prediction models developed using sampling methods is better in terms of accuracy metric as compared to the scenario where no sampling method is used.

An analysis of Table 7 indicates extremely low recall values (5–45%) in majority of the cases on all the data sets when no sampling method is used. A similar trend was shown by precision values as indicated in Table 8. The table shows low precision values (14–65%) in majority of cases when no sampling approach was used. These trends are a result of imbalanced data. As the data sets have very few instances of change prone classes, the models may not be able to learn efficiently the prediction of change prone classes, this would result in low recall and precision values. However, in majority of cases when any sampling approach was used, the recall values ranged from 55 to 98% (Table 7), indicating a positive improvement in the correct prediction of change prone classes. Similarly, the precision values too showed a positive improvement as they ranged from 70 to 96% after applying a sampling approach in majority of cases (Table 8).

The G-mean, balance and AUC performance metrics also showed positive improvement in most of the cases when a sampling method was used as compared to the scenario when no sampling method was used. The G-mean and AUC performance metrics showed a positive improvement of up to 35% with the use of sampling methods and the balance metric of up to 45%.

In order to evaluate whether the use of different sampling methods improve the performance of various ML techniques for ILP we use Friedman test to ascertain the hypothesis given in the following sections. The lower the rank achieved by a given sampling method, the better it is considered. Since this section uses Friedman test to ascertain significant differences amongst four sampling methods namely resample with replacement, SMOTE and spread subsample along with the scenario when no sampling method is used, the degrees of freedom for Friedman test is 3. In order to test the hypothesis, the p-value is stated. The study checks the hypothesis at the cut-off $\alpha = 0.05$. In order to apply Friedman test, we used the values of performance metrics achieved by all the six ML techniques on a specific data set using different sampling approaches and without any sampling approach. This was done for each of the six data sets and the Friedman test was evaluated individually on all the data sets. All the change prediction models were assessed on three stable performance metrics namely G-mean, balance and AUC.

7.4.1 Friedman Results for Different Sampling Methods Using G-mean Measure

This section evaluates hypothesis H0 which is stated below for each data set of the study.

> *Null Hypothesis H0*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) do not show significant differences when various sampling methods (Resampling with replacement, SMOTE, Spread Subsample) are used for handling an imbalanced data set as compared to use of no sampling method when evaluated using G-mean measure.
> *Alternate Hypothesis H0a*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) show significant differences when various sampling methods (Resampling with replacement, SMOTE, Spread Subsample) are used for handling an imbalanced data set as compared to use of no sampling method when evaluated using G-mean measure.

Table 12 states the Friedman ranking allocated to each sampling method and a scenario where no sampling method is used, along with mean ranks allocated by Friedman test (shown in parenthesis) and its p-value, when evaluated using G-mean performance measure on each data set of the study. As depicted in the table, in five out of six data sets, the best rank is achieved by resample method with replacement. However, the worst scenario according to G-mean results was that when no sampling method was used as in five out of six data sets, the no sample scenario achieved the worst rank. The Friedman test was significant at $\alpha = 0.05$ in four data sets (Bluetooth, MMS, Net and Log4j) as a p-value of less than 0.05 was obtained. Since, in majority of the data sets, the resample with replacement method significantly outperformed the other scenarios, we reject null hypothesis H0, which states that use of different sampling methods do not differ significantly from each other when evaluated using G-mean measure.

7.4.2 Friedman Results for Different Sampling Methods Using Balance Measure

This section evaluates Hypothesis H1, which is presented below.

> *Null Hypothesis H1*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) do not show significant differences when various sampling methods (Resampling with replacement, SMOTE, Spread Subsample) are used for handling an imbalanced data set as compared to use of no sampling method when evaluated using balance measure.
> *Alternate Hypothesis H1a*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) show significant differences when various sampling methods (Resampling with replacement, SMOTE, Spread Subsample) are used for handling an imbalanced data set as compared to use of no sampling method when evaluated using balance measure.

Table 12 states the Friedman ranking allocated to each sampling method and a scenario where no sampling method is used, along with mean ranks allocated by Friedman test (shown in parenthesis) and its p-value, when evaluated using balance performance measure on each data set of the study. As depicted in the table, the best rank is achieved by resample method with replacement in all the six data sets. Moreover, the Friedman results were significant at

Table 12 Friedman results using G-mean measure

Data set	Rank 1	Rank 2	Rank 3	Rank 4	p-value
Calendar	Subsample (2.17)	No Sample (2.17)	Resample (2.50)	SMOTE (3.17)	0.494
Bluetooth	Resample (1.17)	Subsample (2.33)	SMOTE (2.50)	No Sample (4.00)	0.002
MMS	Resample (1.33)	SMOTE (1.67)	Subsample (3.17)	No Sample (3.83)	0.002
IO	Resample (2.00)	Subsample (2.33)	SMOTE (2.67)	No Sample (3.00)	0.572
Net	Resample (1.17)	SMOTE (2.17)	Subsample (3.00)	No Sample (3.67)	0.006
Log4j	Resample (1.17)	SMOTE (2.67)	Subsample (2.83)	No Sample (3.33)	0.006

$\alpha = 0.05$ in all the six data sets. The worst scenario according to balance results was that when no sampling method was used as in four out of six data sets the no sample scenario achieved the worst rank. As Resample with replacement method significantly out-performed all other methods in all the six data sets, it leads to acceptance of alternate hypothesis, which means that use of different sampling methods differ significantly from each other when evaluated using balance measure (Table 13).

7.4.3 Friedman Results for Different Sampling Methods Using AUC Measure

This section tests Hypothesis H2 which is stated as follows.

> *Null Hypothesis H2*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) do not show significant differences when various sampling methods (Resampling with replacement, SMOTE, Spread Subsample) are used for handling an imbalanced data set as compared to use of no sampling method when evaluated using AUC measure.
> *Alternate Hypothesis H2a*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) show significant differences when various sampling methods (Resampling with replacement, SMOTE, Spread Subsample) are used for handling an imbalanced data set as compared to use of no sampling method when evaluated using AUC measure.

Table 14 states the Friedman test results when evaluated on all data sets using AUC measure. It reports the ranking allocated to each sampling method and a scenario where no sampling method is used, along with mean ranks allocated by Friedman test (shown in parenthesis) and its p-value. The Friedman results were evaluated at $\alpha = 0.05$. According to Table 14, the resample method with replacement was significantly better than all the other scenarios in all the six data sets. Thus, this leads to rejection of null hypothesis, which means that use of different sampling methods differ significantly from each other when evaluated using AUC measure. Also, in four out of six data sets, the worst rank was allocated to the scenario where no sampling method was used, which indicates that using a sampling approach for an imbalanced data improves the AUC results in majority of the data sets.

Table 13 Friedman results using balance measure

Data set	Rank 1	Rank 2	Rank 3	Rank 4	p-value
Calendar	Resample (1.50)	SMOTE (1.92)	No Sample (3.17)	Subsample (3.42)	0.018
Bluetooth	Resample (1.33)	Subsample (2.17)	SMOTE (2.50)	No Sample (4.00)	0.004
MMS	Resample (1.00)	SMOTE (2.00)	Subsample (3.17)	No Sample (3.83)	0.001
IO	Resample (1.00)	SMOTE (2.00)	Subsample (3.17)	No Sample (4.00)	0.001
Net	Resample (1.00)	SMOTE (2.00)	Subsample (3.17)	No Sample (4.00)	0.001
Log4j	Resample (1.00)	SMOTE (2.00)	No Sample (3.83)	Subsample (3.67)	0.001

Table 14 Friedman Results using AUC measure

Data Set	Rank 1	Rank 2	Rank 3	Rank 4	p-value
Calendar	Resample (1.00)	SMOTE (2.17)	No Sample (3.00)	Subsample (3.83)	0.001
Bluetooth	Resample (1.33)	SMOTE (2.33)	Subsample (2.50)	No Sample (3.83)	0.010
MMS	Resample (1.17)	SMOTE (1.83)	Subsample (3.17)	No Sample (3.83)	0.001
IO	Resample (1.33)	SMOTE (2.00)	Subsample (2.67)	No Sample (4.00)	0.003
Net	Resample (1.00)	SMOTE (1.67)	Subsample (3.00)	No Sample (4.00)	0.001
Log4j	Resample (1.00)	SMOTE (2.17)	No Sample (3.00)	Subsample (3.83)	0.001

7.4.4 Wilcoxon Results for Different Sampling Methods

According to ranks allocated by Friedman test, it was found that the resample with replacement was the best sampling method in majority of the data sets when evaluated using G-mean, balance and AUC performance measures. In order to further investigate and ascertain whether, resample with replacement method is better than all the other employed sampling methods of the study, we use post-hoc Wilcoxon test with Bonferroni correction. The use of Bonferroni correction removes familywise errors. The Wilcoxon test was performed at significance level $\alpha = 0.05$. The test computes pairwise comparisons amongst resample with replacement method and the other sampling methods using G-mean, balance and AUC performance measure. The test was evaluated on the results of all the ML techniques on all the data sets of the study using a specific performance measure. The hypothesis being tested are as follows.

Null Hypothesis H3: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) do not show significant differences in performance metrics (G-mean, balance and AUC) when Resampling with replacement sampling method is used instead of other sampling methods (SMOTE, Spread Subsample) for handling imbalanced data sets.

Alternate Hypothesis H3a: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) show significant differences in performance metrics (G-mean, balance and AUC) when Resampling with replacement sampling method is used instead of other sampling methods (SMOTE, Spread Subsample) for handling imbalanced data sets.

Table 15 Wilcoxon test results on sampling methods performance

Pair	G-mean	Balance	AUC
Resample with replacement vs SMOTE	S+ (0.000)	S+ (0.000)	S+ (0.000)
Resample with replacement vs Spread Subsample	S+ (0.000)	S+ (0.000)	S+ (0.000)

Table 15 shows the results of Wilcoxon test where it is seen that resample with replacement sampling method is significantly better than all the other sampling methods using different performance measures (G-mean, balance, AUC). The table states the p-value of pairwise comparisons of different sampling methods with resample with replacement method. It can be seen that all other sampling methods perform not as well as resample with replacement method and the results of Wilcoxon test are statistically significant. The symbol S+ represents that the results are significantly better according to Wilcoxon test.

RQ2 What is the effectiveness of different MetaCost learners for improving learning through imbalanced data?

In order to ascertain the effectiveness of MetaCost learners, we compare the performance of different MetaCost learners using G-mean, balance, AUC and cost with the scenario when no cost-sensitive approach is applied. In order to calculate the total cost of misclassification, we sum up the misclassification costs of FP and FN predictions as given in equation for each Cost Ratio (CR).

$$\text{Total Cost} = (\text{Cost of a FN Prediction*No. of FN}) + (\text{Cost of an FP prediction*No. of FP})$$

The study analyses seven different CR's which are 5, 10, 15, 20, 25, 30 and 50. Table 16, 17, 18, 19, 20, and 21 states the values of different performance metrics obtained on application of all the six ML techniques (MLP, RF, NB, AB, LB and BG) on each data set of the study using different CR's and the "original" technique without any CR. If a specific performance metric was undefined in a specific scenario, then the value was shown as "N.D." representing "Not Defined". The total cost of each ML method on each data set using various cost ratios is reported in Table 22.

As seen in Table 16, the average performance of all the techniques on all the data sets decreased in terms of accuracy measure when a MetaCost learner was used in comparison to the scenario when no such technique was applied i.e. original ML technique was used. This trend is a result of using a biased performance metric i.e. accuracy which does not take into account the imbalanced nature of the data. Several studies have ascertained the fact that the use of accuracy metric is not suitable as it gives highly optimistic results which are not sensitized to the class distributions in the data set (He and Garcia 2009; Gao et al. 2015; Weng and Poon 2008). The accuracy metric only focuses on the majority class i.e. the large number of non-change prone classes which are correctly predicted and ignores the low number of correctly predicted change prone classes. This fact can be ascertained from the recall values as shown in Table 17. There was a sharp increase in recall values in all the data sets with the use of MetaCost learners. This is because the MetaCost learners were used in such a way so as to penalize the classifier for false negative predictions. The lower, the number of false negative predictions, the higher would be the recall. This was done as the data sets were highly imbalanced with few change prone classes. Thus, it was important to increase recall rates to achieve a balanced classifier. However, one should be careful that apart from recall, the precision values should also be considered. As shown in Table 18, there is decrease in precision values in majority of the cases. This is a concerning factor as increase in recall should not mean that a number of non-change prone classes are incorrectly classified as change prone. One should achieve an optimum balance between recall and precision. It

Table 16 Accuracy results of MetaCost learners using ML techniques

ML tech.	Calendar								ML tech.	Bluetooth							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	75.00	73.00	16.00	17.00	17.00	17.00	17.00	17.00	MLP	83.08	75.38	64.62	56.92	50.77	47.69	44.62	33.85
RF	50.00	68.00	33.00	31.00	29.00	26.00	25.00	23.00	RF	78.46	81.54	64.06	63.08	60.00	58.46	56.92	55.38
NB	41.67	73.00	64.00	34.00	17.00	17.00	17.00	17.00	NB	83.08	84.62	78.46	76.92	75.38	75.38	75.38	75.38
AB	75.00	82.00	17.00	16.00	16.00	17.00	17.00	17.00	AB	75.38	76.92	69.23	67.69	66.15	64.62	61.54	55.38
LB	75.00	78.00	25.00	18.00	16.00	16.00	16.00	17.00	LB	78.46	73.85	72.31	63.08	63.08	55.38	50.77	44.62
BG	50.00	74.00	17.00	17.00	17.00	17.00	17.00	17.00	BG	83.08	76.92	66.15	61.54	46.15	26.15	18.46	16.92
ML tech.	MMS								ML tech.	IO							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	75.72	61.75	25.26	20.00	20.00	20.00	20.00	20.00	MLP	95.72	92.51	80.75	71.66	66.31	64.17	56.15	47.59
RF	71.68	62.11	50.88	45.96	42.81	41.40	39.30	32.98	RF	91.44	86.10	86.10	83.96	82.89	81.28	78.61	75.40
NB	70.52	64.21	31.93	23.16	22.46	22.81	20.70	21.05	NB	93.58	77.54	73.26	72.73	71.12	69.52	67.91	64.71
AB	69.94	56.49	37.54	18.95	20.00	20.00	20.00	20.00	AB	93.05	89.84	89.30	85.03	78.07	77.01	76.47	69.52
LB	73.99	65.26	36.84	22.46	20.00	20.00	20.00	20.00	LB	94.12	91.44	86.10	81.28	73.26	70.05	65.78	44.39
BG	78.61	64.91	32.28	20.00	20.00	20.00	20.00	20.00	BG	94.65	91.44	90.37	86.1	82.89	75.94	62.03	16.04
ML tech.	Net								ML tech.	Log4j							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	76.62	66.67	62.19	52.24	44.78	38.31	38.31	31.34	MLP	81.40	61.75	25.26	20.00	20.00	20.00	20.00	20.00
RF	74.63	68.66	63.68	61.69	59.70	58.71	58.71	60.20	RF	78.60	62.11	50.88	45.96	42.81	41.40	39.30	32.98
NB	74.13	76.12	73.63	66.67	65.17	61.19	61.19	60.70	NB	79.65	64.21	31.93	23.16	22.46	22.81	20.70	21.05
AB	77.11	64.68	56.22	54.73	53.73	52.74	52.74	52.74	AB	82.46	56.49	37.54	18.95	20.00	20.00	20.00	20.00
LB	78.11	63.18	59.70	55.72	55.22	52.24	52.24	31.34	LB	82.81	65.26	36.84	22.46	20.00	20.00	20.00	20.00
BG	77.61	63.18	57.21	55.22	52.74	39.80	39.80	29.35	BG	81.40	64.91	32.28	20.00	20.00	20.00	20.00	20.00

Table 17 Recall results of MetaCost learners using ML techniques

ML tech.	Calendar								ML tech.	Bluetooth							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	17.65	29.41	88.24	100.00	100.00	100.00	100.00	100.00	MLP	27.27	72.73	81.82	81.82	90.91	90.91	90.91	90.91
RF	5.88	41.18	70.59	82.35	82.35	76.47	76.47	76.47	RF	27.27	72.73	80.00	81.82	81.82	81.82	81.82	81.82
NB	29.41	29.41	47.06	70.59	100.00	100.00	100.00	100.00	NB	36.36	72.73	72.73	72.73	72.73	72.73	81.82	90.91
AB	17.65	41.18	94.12	94.12	94.12	100.00	100.00	100.00	AB	9.09	72.73	72.73	72.73	72.73	81.82	81.82	81.82
LB	17.65	41.18	88.24	94.12	94.12	94.12	94.12	100.00	LB	27.27	81.82	72.73	72.73	72.73	72.73	72.73	81.82
BG	17.65	29.41	100.00	100.00	100.00	100.00	100.00	100.00	BG	9.09	72.73	81.82	81.82	81.82	81.82	90.91	100
ML tech.	MMS								ML tech.	IO							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	44.68	63.16	98.25	100.00	100.00	100.00	100.00	100.00	MLP	27.27	27.27	63.64	72.73	81.82	81.82	81.82	81.82
RF	40.43	52.63	63.16	66.67	64.91	66.67	68.42	73.68	RF	27.27	27.27	45.45	45.45	45.45	45.45	45.45	45.45
NB	70.21	50.88	78.95	92.98	92.98	96.49	96.49	98.25	NB	27.27	54.55	81.82	81.82	81.82	81.82	81.82	81.82
AB	51.06	59.65	87.72	94.74	100.00	100.00	100.00	100.00	AB	27.27	36.36	45.45	54.55	54.55	54.55	54.55	63.64
LB	38.30	56.14	85.96	96.49	100.00	100.00	100.00	100.00	LB	27.27	36.36	45.45	45.45	45.45	63.64	63.64	63.64
BG	40.43	43.86	89.47	100.00	100.00	100.00	100.00	100.00	BG	9.09	27.27	45.45	45.45	45.45	45.45	54.55	90.91
ML tech.	Net								ML tech.	Log4j							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	40.68	79.66	94.92	94.92	94.92	100.00	100.00	100.00	MLP	17.54	63.16	98.25	100.00	100.00	100.00	100.00	100.00
RF	50.85	74.58	83.05	84.75	86.44	89.83	89.83	96.61	RF	26.32	52.63	63.16	66.67	64.91	66.67	68.42	73.68
NB	32.20	52.54	72.88	81.36	88.14	91.53	91.53	93.22	NB	14.04	50.88	78.95	92.98	92.98	96.49	96.49	98.25
AB	61.02	83.05	89.83	96.61	96.61	96.61	96.61	98.31	AB	24.56	59.65	87.72	94.74	100.00	100.00	100.00	100.00
LB	57.63	83.05	94.92	96.61	98.31	96.61	96.61	98.31	LB	26.32	56.14	85.96	96.49	100.00	100.00	100.00	100.00
BG	55.93	79.66	86.44	91.53	93.22	100.00	100.00	100.00	BG	19.30	43.86	89.47	100.00	100.00	100.00	100.00	100.00

Table 18 Precision results of MetaCost learners using ML techniques

ML tech.	Calendar								ML tech.	Bluetooth							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	76.39	25.00	15.46	17.00	17.00	17.00	17.00	17.00	MLP	50.00	38.10	30.00	25.71	24.39	23.26	22.22	19.23
RF	80.16	24.14	16.22	17.50	17.07	15.66	15.48	15.12	RF	33.33	47.06	27.59	29.03	27.27	26.47	25.71	25.00
NB	52.70	25.00	22.86	16.44	17.00	17.00	17.00	17.00	NB	50.00	53.33	42.11	40.00	38.10	38.10	39.13	40.00
AB	76.38	46.67	16.33	16.16	16.16	17.00	17.00	17.00	AB	14.29	40.00	32.00	30.77	29.63	30.00	28.13	25.00
LB	76.22	36.84	17.05	16.49	16.16	16.16	16.16	17.00	LB	33.33	37.50	34.78	27.59	27.59	23.53	21.62	20.93
BG	78.21	26.32	17.00	17.00	17.00	17.00	17.00	17.00	BG	50.00	40.00	31.03	28.13	21.43	16.36	16.13	16.92
ML tech.	MMS								ML tech.	IO							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	56.76	56.76	40.91	38.94	39.13	39.13	39.66	38.98	MLP	100.00	33.33	17.95	13.79	12.86	12.16	10.11	8.57
RF	47.50	47.50	41.90	40.00	89.8	39.64	39.29	38.33	RF	27.27	14.29	20.00	17.24	16.13	14.71	12.82	11.11
NB	47.14	47.14	42.86	43.01	41.67	42.86	42.86	45.65	NB	42.86	13.95	15.79	15.52	14.75	14.06	13.43	12.33
AB	45.28	45.28	87.76	39.47	39.13	38.98	38.33	36.22	AB	37.50	25.00	26.32	20.69	14.29	13.64	13.33	11.67
LB	52.94	52.94	40.57	37.29	36.00	35.71	36.51	33.58	LB	50.00	30.77	20.00	14.71	10.20	11.86	10.45	6.54
BG	67.86	67.86	41.12	38.46	38.14	39.50	38.52	32.64	BG	100.00	27.27	29.41	20.00	16.13	11.36	8.33	6.02
ML tech.	Net								ML tech.	Log4j							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	66.67	46.08	43.41	37.58	34.15	32.24	32.24	29.95	MLP	62.50	29.03	20.90	20.00	20.00	20.00	20.00	20.00
RF	57.69	47.83	43.75	42.37	41.13	40.77	40.77	42.22	RF	44.12	27.03	23.23	21.97	20.56	20.43	20.10	19.27
NB	61.29	60.78	53.75	46.15	45.22	42.52	42.52	42.31	NB	47.06	28.16	19.82	19.78	19.63	20.15	19.71	20.00
AB	61.02	44.55	39.26	39.04	38.51	38.00	38.00	38.16	AB	66.67	25.19	22.62	19.15	20.00	20.000	20.00	20.00
LB	64.15	43.36	41.79	39.58	39.46	37.75	37.75	29.74	LB	68.18	30.19	22.17	20.07	20.00	20.00	20.00	20.00
BG	63.46	43.12	39.53	38.85	37.67	32.78	32.78	29.35	BG	61.11	26.88	21.43	20.00	20.00	20.00	20.00	20.00

Table 19 G-Mean results of MetaCost learners using ML techniques

ML tech.	Calendar								ML tech.	Bluetooth							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	0.80	0.46	0.23	N.D.	N.D.	N.D.	N.D.	N.D.	MLP	0.66	0.60	0.53	0.49	0.48	0.47	0.46	0.42
RF	0.65	0.46	0.36	0.39	0.38	0.35	0.34	0.33	RF	0.53	0.66	0.51	0.52	0.50	0.49	0.48	0.48
NB	0.59	0.46	0.44	0.36	N.D.	N.D.	N.D.	N.D.	NB	0.66	0.70	0.63	0.61	0.60	0.59	0.61	0.62
AB	0.80	0.64	0.29	0.00	0.00	N.D.	N.D.	N.D.	AB	0.34	0.61	0.54	0.53	0.52	0.53	0.51	0.48
LB	0.80	0.57	0.38	0.33	0.00	0.00	0.00	N.D.	LB	0.53	0.59	0.57	0.50	0.50	0.46	0.44	0.44
BG	0.65	0.47	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	BG	0.65	0.61	0.54	0.51	0.44	0.36	0.33	N.D.
ML tech.	MMS								ML tech.	IO							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	0.68	0.68	0.63	0.61	0.61	0.61	0.62	0.62	MLP	0.98	0.56	0.42	0.37	0.36	0.35	0.31	0.29
RF	0.61	0.61	0.63	0.62	0.93	0.61	0.61	0.61	RF	0.51	0.37	0.44	0.41	0.39	0.38	0.35	0.33
NB	0.64	0.61	0.62	0.63	0.62	0.63	0.63	0.65	NB	0.64	0.37	0.39	0.39	0.38	0.37	0.36	0.34
AB	0.61	0.61	0.91	0.62	0.61	0.62	0.6	0.60	AB	0.59	0.49	0.50	0.45	0.37	0.36	0.36	0.34
LB	0.65	0.65	0.62	0.59	0.59	0.58	0.60	0.57	LB	0.69	0.54	0.44	0.38	0.31	0.34	0.32	0.25
BG	0.74	0.74	0.63	0.61	0.61	0.63	0.62	0.57	BG	0.97	0.51	0.53	0.44	0.39	0.33	0.28	0.24
ML tech.	Net								ML tech.	Log4j							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	0.72	0.64	0.64	0.59	0.56	0.57	0.57	0.55	MLP	0.79	0.50	0.44	N.D.	N.D.	N.D.	N.D.	N.D.
RF	0.68	0.64	0.62	0.61	0.61	0.61	0.61	0.64	RF	0.89	0.48	0.44	0.30	0.41	0.41	0.40	0.39
NB	0.68	0.70	0.68	0.64	0.64	0.63	0.63	0.63	NB	0.74	0.49	0.40	0.39	0.38	0.41	0.36	0.40
AB	0.72	0.63	0.60	0.61	0.61	0.60	0.60	0.61	AB	0.76	0.46	0.45	0.00	N.D.	N.D.	N.D.	N.D.
LB	0.73	0.62	0.63	0.62	0.62	0.60	0.60	0.50	LB	0.81	0.51	0.44	0.41	N.D.	N.D.	N.D.	N.D.
BG	0.72	0.61	0.59	0.60	0.59	0.57	0.57	N.D.	BG	0.85	0.47	0.43	N.D.	N.D.	N.D.	N.D.	N.D.

Table 20 Balance results of MetaCost learners using ML techniques

ML tech.	Calendar								ML tech.	Bluetooth							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	41.76	48.48	29.65	29.29	29.29	29.29	29.29	29.29	MLP	48.42	74.28	69.64	63.61	58.9	56.31	53.72	44.63
RF	33.44	54.38	43.23	42.40	40.74	38.09	37.27	35.62	RF	47.98	77.40	69.08	68.45	66.04	64.83	63.61	62.38
NB	49.73	48.48	56.06	44.02	29.29	29.29	29.29	29.29	NB	54.70	78.65	75.93	75.12	74.28	74.28	77.61	79.33
AB	41.76	57.85	30.02	29.17	29.17	29.29	29.29	29.29	AB	35.24	75.12	70.55	69.55	68.52	69.64	67.25	62.38
LB	41.76	57.17	37.25	30.87	29.17	29.17	29.17	29.29	LB	47.98	76.52	72.47	66.41	66.41	60.87	57.41	53.66
BG	41.71	48.68	29.29	29.29	29.29	29.29	29.29	29.29	BG	35.70	75.12	70.83	67.25	54.92	38.41	31.61	29.29
ML tech.	MMS								ML tech.	IO							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	59.87	59.87	63.40	61.02	60.60	60.60	60.69	59.57	MLP	48.57	48.52	71.25	72.15	72.32	70.89	65.38	59.34
RF	56.26	56.26	65.47	62.69	93.00	62.13	61.57	58.44	RF	48.47	48.07	60.6	60.24	60.04	59.71	59.08	58.22
NB	70.42	63.7	68.43	68.45	66.86	67.69	67.69	69.59	NB	48.55	64.59	76.82	76.49	75.47	74.43	73.39	71.25
AB	61.76	61.76	91.33	61.16	60.60	59.57	58.44	54.52	AB	48.53	54.74	61.02	66.56	64.75	64.42	64.24	66.61
LB	55.46	55.46	64.14	58.23	55.00	54.44	52.84	48.91	LB	48.56	54.86	60.6	59.71	57.57	66.87	64.75	52.30
BG	57.57	57.57	64.36	59.48	58.92	59.59	57.91	45.56	BG	35.72	48.47	61.13	60.60	60.04	58.37	58.33	37.00
ML tech.	Net								ML tech.	Log4j							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	57.63	69.07	63.47	53.55	46.10	38.25	38.25	31.28	MLP	41.67	62.27	34.24	29.29	29.29	29.29	29.29	29.29
RF	63.56	70.09	66.42	64.46	62.41	60.99	60.99	61.09	RF	47.57	58.13	54.83	51.95	49.18	48.40	47.00	42.33
NB	51.69	65.00	73.41	69.15	67.53	63.16	63.16	62.35	NB	39.15	58.37	41.63	33.14	32.52	32.35	30.49	30.52
AB	70.15	67.35	58.54	55.62	54.62	53.63	53.63	53.18	AB	46.61	57.63	46.26	29.19	29.29	29.29	29.29	29.29
LB	68.58	65.95	60.99	56.61	55.67	53.13	53.13	31.77	LB	47.85	61.42	45.74	32.04	29.29	29.29	29.29	29.29
BG	67.43	65.94	59.99	57.25	54.43	39.75	39.75	29.29	BG	42.89	55.05	41.53	29.29	29.29	29.29	29.29	29.29

Table 21 AUC results of MetaCost learners using ML techniques

ML tech.	Calendar								ML tech.	Bluetooth							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	0.53	0.54	0.41	0.42	0.42	0.42	0.42	0.42	MLP	0.79	0.81	0.82	0.81	0.76	0.78	0.73	0.66
RF	0.61	0.52	0.44	0.47	0.48	0.49	0.49	0.48	RF	0.70	0.78	0.76	0.80	0.79	0.77	0.77	0.76
NB	0.58	0.62	0.61	0.54	0.56	0.59	0.59	0.59	NB	0.85	0.87	0.86	0.87	0.86	0.86	0.86	0.87
AB	0.43	0.66	0.44	0.47	0.48	0.5	0.50	0.50	AB	0.66	0.76	0.70	0.74	0.72	0.73	0.72	0.69
LB	0.49	0.61	0.43	0.46	0.39	0.5	0.50	0.51	LB	0.67	0.78	0.73	0.76	0.75	0.66	0.69	0.69
BG	0.54	0.56	0.50	0.50	0.50	0.50	0.50	0.50	BG	0.77	0.77	0.78	0.73	0.63	0.48	0.50	0.50
ML tech.	MMS								ML tech.	IO							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	0.80	0.54	0.70	0.71	0.71	0.71	0.72	0.75	MLP	0.74	0.68	0.69	0.79	0.76	0.79	0.74	0.73
RF	0.61	0.52	0.78	0.73	0.71	0.71	0.72	0.74	RF	0.67	0.64	0.68	0.70	0.68	0.67	0.68	0.69
NB	0.79	0.62	0.77	0.78	0.78	0.78	0.78	0.78	NB	0.72	0.76	0.77	0.77	0.77	0.77	0.76	0.74
AB	0.76	0.66	0.79	0.79	0.80	0.80	0.80	0.80	AB	0.66	0.63	0.68	0.71	0.75	0.75	0.77	0.76
LB	0.79	0.61	0.80	0.76	0.78	0.78	0.81	0.78	LB	0.70	0.68	0.74	0.73	0.75	0.73	0.74	0.70
BG	0.79	0.56	0.71	0.72	0.70	0.70	0.74	0.61	BG	0.67	0.71	0.71	0.66	0.70	0.70	0.58	0.51
ML tech.	Net								ML tech.	Log4j							
	Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50		Original	MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	0.77	0.76	0.74	0.75	0.75	0.53	0.51	0.30	MLP	0.67	0.70	0.52	0.36	0.31	0.30	0.29	0.29
RF	0.80	0.79	0.78	0.75	0.73	0.73	0.72	0.72	RF	0.61	0.58	0.53	0.52	0.51	0.51	0.57	0.49
NB	0.78	0.78	0.79	0.79	0.78	0.79	0.77	0.75	NB	0.62	0.61	0.54	0.60	0.60	0.60	0.61	0.62
AB	0.78	0.79	0.79	0.71	0.70	0.72	0.72	0.76	AB	0.66	0.62	0.65	0.50	0.50	0.50	0.50	0.50
LB	0.79	0.78	0.80	0.79	0.77	0.78	0.78	0.50	LB	0.68	0.65	0.66	0.60	0.52	0.52	0.52	0.52
BG	0.78	0.76	0.69	0.66	0.68	0.63	0.54	0.50	BG	0.68	0.60	0.60	0.50	0.50	0.50	0.50	0.50

Table 22 Cost values of MetaCost learners using ML techniques

ML tech.	Calendar							ML tech.	Bluetooth						
	MC5	MC10	MC15	MC20	MC25	MC30	MC50		MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	75	102	83	83	83	83	83	MLP	28	41	56	51	58	65	92
RF	72	112	111	128	170	191	273	RF	24	41	52	64	75	86	127
NB	75	117	136	83	83	83	83	NB	22	41	57	73	88	74	65
AB	58	92	98	103	83	83	83	AB	27	47	63	79	71	83	127
LB	62	93	96	103	108	113	83	LB	25	45	66	81	101	119	134
BG	74	83	83	83	83	83	83	BG	27	40	53	73	96	82	54
ML tech.	MMS							ML tech.	IO						
	MC5	MC10	MC15	MC20	MC25	MC30	MC50		MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	146	85	114	110	120	100	122	MLP	46	72	95	101	115	140	196
RF	161	91	111	65	142	158	124	RF	58	80	114	146	179	214	340
NB	157	132	158	196	181	206	300	NB	62	68	79	92	105	118	164
AB	144	46	99	110	97	104	131	AB	47	74	98	136	163	189	253
LB	161	103	119	120	131	110	141	LB	44	80	119	164	152	180	300
BG	149	93	102	113	72	75	97	BG	48	72	110	146	189	216	206
ML tech.	Net							ML tech.	Log4j						
	MC5	MC10	MC15	MC20	MC25	MC30	MC50		MC5	MC10	MC15	MC20	MC25	MC30	MC50
MLP	115	103	138	168	124	124	138	MLP	193	222	228	228	228	228	228
RF	123	163	203	233	227	257	178	RF	216	329	420	543	623	695	926
NB	160	197	221	203	198	223	275	NB	214	302	275	297	268	284	274
AB	111	142	119	131	143	153	144	AB	216	241	273	228	228	228	228
LB	114	108	117	109	144	154	187	LB	199	252	249	228	228	228	228
BG	122	158	160	171	121	121	142	BG	228	247	228	228	228	228	228

Table 23 Wilcoxon test results on resample method vs MetaCost learners

Data set	G-Mean	Balance	AUC
Calendar	S+ (0.028)	S+ (0.046)	S+ (0.046)
Bluetooth	S+ (0.028)	S+ (0.046)	S+ (0.027)
MMS	S+ (0.046)	S+ (0.075)	S+ (0.027)
IO	S+ (0.027)	S+ (0.075)	S+ (0.027)
Net	S+ (0.027)	S+ (0.028)	S+ (0.027)
Log4j	S+ (0.027)	S+ (0.046)	S+ (0.028)

should be a researcher's objective to increase both recall as well as precision and both change-prone and not change-prone classes should be correctly identified. For example, result of CR = 5 in Bluetooth and MMS data sets show better or comparable precision values.

After performing the above analysis, it can be seen that the performance of different ML techniques varied significantly using different sampling methods when evaluated using G-mean, balance or AUC performance measure. The mean ranks obtained by the above Friedman tests advocate the resample method with replacement as the best sampling method in majority of the cases. These results of Friedman tests are further supported by the use of Wilcoxon test, according to which the resample with replacement method is significantly better than the other two sampling methods when pairwise comparisons are performed. Furthermore, it was analysed that the performance of different sampling methods was better than the scenario where no sampling method was used.

Table 19 shows the G-mean values with different CR's on all the data sets of the study. It should be noted that there was decrease in G-mean values in four data sets and improvement in G-mean values in only two data sets (Bluetooth and MMS). This is because G-mean is the geometric mean of both accuracies of change prone as well as non-change prone classes. MetaCost learners tend to increase the recall rates i.e. the accuracy of predicting change prone classes which can lead to decrease in the accuracy with which non-change prone classes are predicted. However, this decrease in accurate prediction of non-change prone classes may not be compensated well by the increase in recall values as change prone classes are fewer in number. The decrease in precision values also indicates this trend. Thus, leading to decrease in overall G-mean values for four data sets of the study.

Table 24 Data set details for inter-release validation

Software name	Versions for training	Versions for validation
Android calendar	4.0.2–4.0.4 (35%)	4.0.4–4.1.2 (61%)
Android bluetooth	4.3.1–4.4.2 (16%)	5.0.2–5.1.0 (30%)
Android MMS	2.3.7–4.0.2 (27%)	4.1.2–4.2.2 (52%)
Apache IO	1.3–1.4 (30%)	1.4–2.0 (60%)
Apache net	3.0–3.1 (29%)	3.1–3.2 (51%)
Apache Log4j	1.2.13–1.2.14 (6%)	1.2.16–1.2.17 (25%)

Table 20 shows the balance results of different MetaCost learners on each data set of the study. It was noted that various CR's gave an improved balance results on different data sets. The balance results on Calendar data set was improved up to a maximum of 11%. Similarly, the balance results on Bluetooth, MMS, IO, Net and Log4j data sets were improved up to a maximum of 31, 50, 20, 4 and 15% respectively. Table 21 reports the AUC results of different MetaCost learners. It was noted that certain CR's gave a maximum improvement of 5, 5, 23, 5 and 2% correspondingly in the AUC values of Calendar, Bluetooth, MMS, IO and Log4j data sets respectively. These results favour the use of MetaCost learners as compared to the "original" technique.

We also applied Friedman test on G-mean, balance and AUC values of each data set individually to rank the capability of different CR's along with the "original" technique to evaluate whether the improvement in different performance metrics is statistically significant or not when MetaCost learners are used. The Friedman test was conducted at significance level $\alpha = 0.05$. When Friedman test was conducted on AUC values, four out of six data sets showed that MetaCost learners gave better value than the "original" technique. Out of these four data sets, three gave statistically significant results indicating that MetaCost learners showed statistically significant improvement than the "original" technique. The application of Friedman test on balance values indicated that in five out of six data sets, the MetaCost learners statistically outperformed the "original" technique. However, in only two out of six data sets, the MetaCost learners statistically outperformed the "original" technique on G-mean values. As discussed previously, this decrease in G-mean values is due to decrease in accurate prediction of non-change prone classes which is not compensated by correct prediction of change prone classes as change prone classes are few in number. However, since in majority of cases MetaCost learners gave statistically improved results than the "original" technique, when evaluated using G-mean, balance and AUC performance metrics, they are favourable for learning through imbalanced data.

Upon evaluating MetaCost learners with different CR's, each data set gave a best model with correspondingly different CR. Also, the cost of the model for each CR ratio (shown in Table 22) was considered while choosing the best CR ratio for a specific data set. The Android Bluetooth and Calendar data sets, along with Apache Net and Log4j data sets gave the best MetaCost model with CR = 5. However, the Android MMS data set gave the best performing MetaCost learner with CR = 10 and Apache IO with CR = 15. Thus, we advocate that developers should explore different CR's in order to find the best MetaCost learners. A similar recommendation is given by Seliya and Khoshgoftaar (2011).

Thus, after analysing the results produced by different MetaCost learners, we can see that MetaCost learning is an efficient way to handle imbalanced learning problem. In general, in majority of the cases the results of balance and AUC performance metrics improved statistically as compared to the original learning technique in all the data sets. However, it may be noted that use of different CR's may provide different results. A researcher needs to assess different CR's for a specific data set in order to ascertain the CR which provides the best model.

RQ3 What is the comparative performance of the best sampling method and MetaCost learner for ILP?

According to RQ1, the best sampling method was resample with replacement. This RQ evaluates the predictive performance of change prediction models developed using resample with replacement method with the best MetaCost learner for each data set.

The performance evaluation is done on the basis of G-mean, balance and AUC values of change prediction models by performing Wilcoxon test at the cut-off $\alpha = 0.05$. The test was performed on each data set individually. We tested the following hypothesis H4.

> *Null Hypothesis H4*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) do not show significant differences in performance metrics (G-mean, balance and AUC) when Resampling with replacement sampling method is used instead of MetaCost learners.
> *Alternate Hypothesis H4a*: Change prediction models developed using various ML Techniques (MLP, RF, NB, AB, LB, BG) show significant differences in performance metrics (G-mean, balance and AUC) when Resampling with replacement sampling method is used instead of MetaCost learners.

In order to apply Wilcoxon test, we used G-mean, balance and AUC values of all the ML techniques on each data set respectively. Table 23 reports the results of Wilcoxon test and its p-value (shown in parenthesis) when the performance metrics obtained by Resample method are compared with the performance metrics using MetaCost learners. According to the results, the Resample with replacement method significantly out-performed the MetaCost learners in majority of the cases (p-value < 0.05). S+ indicates the significantly better performance of the Resample with replacement method in the specific case. Only in MMS and IO data sets balance values, the results are not significant, though in these two cases too the resample method showed better results than the corresponding MetaCost learner. Thus, we accept the alternate hypothesis, which means that the Resample with replacement method is significantly better than the MetaCost learners.

On comparative analysis of the best sampling method i.e. resample with replacement with MetaCost learners, it was found that the resample with replacement method develops better change prediction models. The performance evaluation was done on the basis of G-mean, balance and AUC performance metrics. The results were statistically analysed using Wilcoxon test. The test supported the resample with replacement technique over MetaCost learners as the sampling method yielded statistically significantly better results than MetaCost learners in majority of the cases on each of the six data sets of the study.

8 Inter-release Validation of Change Prediction Models

This section evaluates the results of inter-release prediction for models which are developed using imbalanced data sets. In order to do so, we first developed training models using a specific release of a data set which is imbalanced in nature. Thereafter, these developed models were validated on a separate release of the corresponding data set. Table 24 reports the data set versions used for training and model development as well as the versions used for model validation. The table also denotes the percentage of changed classes in each data set (shown in parenthesis). The steps for developing inter-release validation models are similar to those for developing ten-fold cross validation models. We first remove outliers from each training data set. It may be noted that the reported percentage of change in training data set is the one which we get after eliminating these outliers. The next step involves application of CFS for reducing the number of independent variables in the training data. After this step, we apply sampling methods or MetaCost learners for developing training models. These training models are

then validated using another release of the same data set, which is specified in Table 24. The validation data will have only those independent variables which were selected by CFS in the corresponding training data set. However, we do not perform any outlier removal in validation data sets.

We now evaluate the results of inter-release validation and discuss the answers to all the three RQ's in light of these results.

> RQ1a: Does the performance of different ML methods in this study significantly improve by using various sampling approaches for ILP in change prediction?
> RQ1b: Which sampling approach performs best amongst different sampling techniques analysed in the study?

After analysing the accuracy results of inter-release validation on various data sets of the study, it was found that in majority of the cases in four data sets (Calendar, MMS, Net and IO), the accuracy values showed an improvement after application of sampling techniques. However, in two data sets (Log4j and Bluetooth), the accuracy values declined but as discussed earlier, accuracy is not an effective indicator for imbalanced data sets (He and Garcia 2009; Gao et al. 2015; Weng and Poon 2008). The recall values showed positive improvement after application of sampling techniques on all the data sets for inter-release validation results as more number of change prone classes were correctly predicted after application of sampling techniques. In four out of six data sets, the precision results decreased after application of sampling techniques. However, in majority of cases, there was only a decrease of 10% in precision values but there was an increase of 30% in recall values in certain cases. Since, the precision values are generally assessed together with recall, the decrease in precision values is not much as compared to improvement of recall values. As discussed by Menzies et al. (2007b), though predictive models should attain both high precision and recall values but in order to optimize one of these measures, the other may have to be compromised especially for imbalanced data sets with low number of positive (change prone) instances. Also, a certain decrease in precision values means that few resources might be wasted in testing classes which are not change prone but are predicted as change prone but critical change prone classes will not be missed by software practitioners as the models exhibit high recall values.

The G-mean results using inter-release validation showed positive improvement when sampling techniques were used in four data sets. However, these results were significant in only two data sets (Net and Log4j), when evaluated using the Friedman test at $\alpha = 0.05$. The balance and AUC results also improved in majority of the data sets when different sampling approaches were applied for handling imbalanced data sets. Four (Calendar, Bluetooth, MMS and Net) out of six data sets showed significant improvement in balance results when Friedman test was applied on balance values of different data sets individually. The other two data sets also showed improved balance results but not significantly. The AUC results obtained using inter-release validation showed positive improvement in five data sets, when a sampling method was used for model development. However, only the IO data set showed a significant improvement in AUC results using Friedman test. Thus, we conclude that the application of sampling methods improve the performance of different ML techniques for handling ILP, when evaluated using G-mean, balance and AUC values. However, these results were significantly better for various

sampling approaches in only 7 out of 18 scenarios as Friedman test was individually applied on each of the six data sets of the study for G-mean, balance and AUC values.

In order to evaluate which sampling method performs best, we analysed the seven significant cases of Friedman test application on G-mean, balance and AUC values of different data sets individually. In four out of these seven significant cases, the resample with replacement method achieved the best rank advocating its better performance as compared to other sampling methods. Thus, our inter-release validation results are comparable to ten-fold cross validation results as Resample with replacement method is advocated as the best sampling method for handling ILP.

> RQ2 What is the effectiveness of different MetaCost learners for improving learning through imbalanced data?

The inter-release validation results were computed on seven different CR's, which were 5, 10, 15, 20, 25, 30 and 50, similar to the manner in which ten-fold cross validation results were computed. These results were compared with the "original" scenario when no cost-sensitive approach was applied. On evaluating the inter-release validation results, we found that the accuracy values showed improvement in only three data sets (Net, IO and MMS). However, in five out of six data sets, the recall values improved drastically. Also, in three (Net, Calendar and IO) out of six data sets, the precision values were improved or remained same. However, there was a decrease in the precision values in the other three data sets. As we discussed earlier, accuracy is not an effective indicator for ILP due to its biased nature and precision values may have to be compromised to achieve high recall values (He and Garcia 2009; Gao et al. 2015; Menzies et al. 2007b; Weng and Poon 2008). Thus, these trends are acceptable.

In most of the data sets, there was an increase in G-mean, balance and AUC values when MetaCost learners were used as compared to the "original" technique, when no MetaCost learner was used. We applied Friedman test at $\alpha = 0.05$ on the values obtained by all ML methods on each of the six data sets of the study. The Friedman test was individually applied on G-mean, balance and AUC results of inter-release validation on each data set of the study. In five out of six data sets, the balance values obtained after using specific CR's significantly outperformed the balance values of the "original" technique. In four out of six data sets, the AUC values obtained using MetaCost learners achieved better Friedman ranks as compared to the "original" techniques. However, the Freidman results were significant in only three data sets when compared using AUC values. When evaluating the G-mean results, it was found that in two data sets the G-mean values obtained using the MetaCost learners significantly outperformed the "original" technique. In all, out of 18 scenarios in which Friedman test was applied, there were 13 cases, when Friedman results were significant. Out of these 13 cases, in nine scenarios, the performance metrics obtained by MetaCost learners significantly outperformed the performance metrics obtained by the "original" technique. Thus, these results advocate the effectiveness of MetaCost learners for improving learning through imbalanced data.

Also, it may be noted that the CR which gave the best change prediction model differed for each specific data set. The Log4j and Calendar data sets gave the best inter-release change prediction models with CR = 5. The best performing MetaCost learner for

Bluetooth and IO data sets was CR = 10. However, the MMS data set obtained the best inter-release change prediction results with CR = 15 and for Net data set, it was CR = 25.

> RQ3 What is the comparative performance of the best sampling method and MetaCost learner for ILP?

As indicated in RQ1, the performance of various sampling methods when evaluated using inter-release validation yielded resample with replacement as the best method. In this RQ, we compare the inter-release change prediction models developed using resample with replacement method with the inter-release change prediction models developed using the best MetaCost learner on a specific data set. The comparison was done according to the G-mean, balance and AUC values of change prediction models on each of the six data sets of the study using Wilcoxon test at $\alpha = 0.05$.

The pairwise comparison results of resample method with MetaCost learners on all the six data sets using G-mean values indicated that the performance of MetaCost learners was significantly better than the resample method in two data sets (MMS and Net). However, in Bluetooth data set the G-mean values of resample method were significantly better than the MetaCost learner. In all the other three data sets, though the performance of resample method was better than the MetaCost learners, the results were not significant.

The Wilcoxon test results on the basis of balance values of inter-release change prediction models indicated that in four out of six data sets, the Resample method performed better than the MetaCost learner. However, these results were not significant. The MetaCost learners outperformed the resample method in the other two data sets using balance values, but again these results were not significant.

The results of Wilcoxon test on the AUC values of inter-release change prediction models indicated that the resample method outperformed the MetaCost learners on five data sets of the study. However, the results were significant in only one case. Moreover, in Net data set, the results of resample method and MetaCost learner was equivalent.

The above discussed results indicate that though the Resample with replacement method is better than the MetaCost learners in majority of the cases, the results are not significant. Thus, we conclude that the results of inter-release change prediction models are comparable to that of ten-fold cross validation models as the Resample with replacement method could not significantly outperform the MetaCost learners using G-mean, balance and AUC values of inter-release change prediction models.

9 Threats to Validity

The various threats to validity of the study are discussed in this section.

9.1 Conclusion Validity

Threats to conclusion validity includes sources which may lead to misjudgement of accurate relationship between the dependent and the independent variables (Catal and Diri 2009; Khoshgoftaar et al. 2006). It is also referred to as statistical conclusion

validity. We have been very careful in selection of appropriate statistical tests for analysing the results of the studies. Since the study uses non-parametric tests, they do not require normality pre-conditions to be checked on the underlying data. Moreover, the study uses a widely accepted α value of 0.05 to check proper significance of results. The use of post-hoc tests further increase the confidence in the results of the study. The use of multiple performance evaluators is effective in strengthening the conclusions of the study. Moreover, the use of ten-fold cross validation and ten iterations in the case of resample with replacement method aids in eliminating validation bias of the study.

9.2 Internal Validity

The threat to internal validity associates itself with the "causal effect" of the independent variables on the dependent variable i.e. the change proneness attribute of a class (Zhou et al. 2009). This effect can only be evaluated with the aid of controlled experiments, which are very challenging to actually perform (Briand et al. 2001). Since, evaluation of "causal effect" was not the primary aim of the study, we did not evaluate this effect. Thus, this threat exists in our study. Also, the outliers have been removed for both training as well as validation data, when ten-fold cross validation method have been used for model development. This could also be a possible source of threat to the results of the study. However, since we also ascertain the results on inter-release validation models, where outliers have been only removed from the training data and the results obtained using inter-release validation models are comparable to ten-fold cross validation results, thus this threat is minimized. The study uses ten-fold cross validation with resampling method, where the resampled folds are used for evaluation. This is probable source of threat to the study. However, in inter-release validation, the evaluation is done on different data sets and not resampled folds and the results obtained are comparable to that of ten-fold cross validation. Thus, this threat is also reduced.

9.3 Construct Validity

Construct validity is concerned with the degree to which the independent and dependent variables accurately measure the concepts they represent (Zhou et al. 2009). Since, the study uses the LOC metric, extracted from the change logs using the DCRS tool to evaluate the dependent variable, we are assured of its accuracy. Certain researchers have tried to evaluate the accuracy of various OO metrics selected in the study, which assures us of their correctness (Briand et al. 1998; Briand et al. 1999; Briand et al. 2000). Moreover, all the OO metrics used in the study are widely accepted and used in the software engineering community. Thus, the threat to construct validity is minimized.

9.4 External Validity

The external validity is concerned with the extent to which the results of the study are generalizable. This study uses six popular open source data sets which are developed in Java language for empirical validation. Thus, the results of the study can be widely used

in similar scenarios. Moreover, the study clearly states the parameter settings of all the ML techniques used, which aids the replicability of results.

10 Conclusions and Future Directions

The study developed and analysed change prediction models using six imbalanced software data sets. In order to deal with imbalanced learning problem, the study used three data sampling methods (Resample with replacement, Spread subsample and SMOTE) and MetaCost learners using seven different cost ratios. The change prediction models were developed using six ML techniques. The developed models were validated using two approaches: ten-fold cross validation and inter-release validation and were assessed using three unbiased and robust performance metrics namely G-mean, Balance and AUC. The trend of traditional metrics like accuracy, recall and precision was also observed on models developed by using different techniques for imbalanced learning. The main contribution of the study is to advocate the use of different methods for handling imbalanced data sets as use of such methods leads to vast improvement in results of the change prediction models. The study also involved statistical analysis of the results produced in order to strengthen the conclusions of the study. The conclusions of the study are reported as follows:

- The performance of various ML techniques significantly improved after use of different sampling methods. Moreover, the resample method with replacement gave the best results when evaluated in terms of G-mean, balance and AUC performance metrics. However, the other two sampling methods i.e. SMOTE and spread subsample also showed good results.
- The use of MetaCost learners is yet another efficient way of handling imbalanced learning problem as they sensitize the ML method by providing different costs for different kinds of misclassification errors so that the total overall cost of misclassification is decreased and the performance of the developed change prediction model improves. However, one should evaluate different cost ratios on a specific data set to identify the best cost ratio that works well on a specific data set.
- A comparative performance of the best sampling method i.e. resample with replacement with the MetaCost learner having the most effective cost ratio on all the data sets reveal that the sampling method significantly outperforms the MetaCost learners when evaluated on the basis of G-mean, balance and AUC performance metrics on ten-fold cross validation models. The results were supported by inter-release validation models. This result advocates that providing effective training data using resampling is a better approach for handling imbalanced learning rather than cost-sensitizing the learners.

Thus, the results of the study can be used by software practitioners and researchers to develop effective change prediction models when faced with imbalanced data sets. Moreover, such models can be built using optimum costs. Future work may include replicated studies with data sets from other domains and programming languages. Furthermore, apart from ML techniques, we may evaluate whether the discussed

approaches for handling imbalanced data sets are effective when search-based techniques are used for model development.

References

Apandi ZFM, Mustapha N, Affendey LS (2011) Evaluating integrated weight linear method to class imbalanced learning in video data. In 3rd Conference on *Data Mining and Optimization,* 243–247

Arisholm E, Briand LC, Johannessen EB (2010) A systematic and comprehensive investigation of methods to build and evaluate fault prediction models. J Syst Softw 83(1):2–17

Bekkar M, Djemaa HK, Alitouche TA (2013) Evaluation measures for models assessment over imbalanced data sets. J Inf Eng Appl 3(10):27–38

Bieman J, Jain D, Yang H (2001) OO design patterns, design structure, and program changes: an industrial case study. In proceedings of 17th International Conference on Software Maintenance, 580–589

Breiman L (1996) Bagging predictors. Mach Learn 24:123–140

Briand L, Daly J, Wust J (1998) A unified framework for cohesion measurement in object-oriented systems. Empir Softw Eng 3(1):65–117

Briand L, Daly J, Wust J (1999) A unified framework for coupling measurement in object-oriented systems. IEEE Trans Softw Eng 25(1):91–121

Briand L, Wust J, Daly JW (2000) Exploring the relationship between design measures and software quality in object-oriented systems. J Syst Softw 51(3):245–273

Briand L, Wust J, Lounis H (2001) Replicated case studies for investigating quality factors in object oriented designs. Empir Softw Eng J 6(1):11–58

Brieman L (2001) Random forests. Mach Learn 45(1):5–32

CartWright M, Shepperd M (2000) An empirical investigation of an object-oriented software system. IEEE Tran Softw Eng 26(8):786–796

Carvalho ABD, Pozo A, Vergilio SR (2010) A symbolic fault-prediction model based on multi-objective particle swarm optimization. J Syst Softw 83(5):868–882

Catal C, Diri B (2009) Investigating the effect of dataset size, metrics sets, and feature selection techniques on software fault prediction problem. Inf Sci 179(8):1040–1058

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357

Chidamber SR, Kemerer CF (1994) A metrics suite for object oriented design. IEEE Tran Softw Eng 20(6):476–493

Demšar J (2006) Statistical comparisons of classifiers over multiple data sets. J Mach Learn Res 7:1–30

Domingos P (1999) Metacost: a general method for making classifiers cost-sensitive. In Proc. of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, CA, 155–164

Elish MO, Al-Khiaty MA (2013) A suite of metrics for quantifying historical changes to predict future change-prone classes in object-oriented software. J Softw: Evol Process 25(5):407–437

Eski S, Buzluca F (2011) An empirical study on object-oriented metrics and software evolution in order to reduce testing cost by predicting change prone classes. In Proc. of International Conference on Software Testing, Verification and Validation Workshop, 566–571

Fawcett T (2006) An introduction to ROC analysis. Pattern Recogn Lett 27(8):861–874

Friedman J, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting. Ann Stat 28(2):337–407

Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F (2012) A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern Part C Appl Rev 42(4):463–484

Gao K, Khoshgoftaar TM, Napolitano A (2015) Combining feature subset selection and data sampling for coping with highly imbalanced software data. In Proc. of 27th International Conf. on Software Engineering and Knowledge Engineering, Pittsburgh, 2015

Giger E, Pinzger M, Gall HC (2012) Can we predict type of code changes? An empirical analysis. In Proc. of 9th IEEE Working Conference on Mining Software Repositories, 217–226

Hall MA (2000) Correlation-based feature selection for discrete and numeric class machine learning. In Proc. of the Seventeenth International Conference on Machine Learning, 359–366

Hall MA, Holmes G (2003) Benchmarking attribute selection techniques for discrete class data mining. IEEE Trans Knowl Data Eng 15(6):1437–1447

Harman M, Islam S, Jia Y, Minku LL, Sarro F, Sirivisut K (2014) less is more: temporal fault predictive performance over multiple Hadoop releases. In Proc. 6th International Symposium on Search Based Software Engineering, 240–246

Haykin S (2004) Neural networks: a comprehensive foundation, 2nd edn. Pearson education, Delhi

He H, Garcia EA (2009) Learning from imbalanced data. IEEE Trans Knowl Data Eng 21(9):1263–1284

Henderson-Sellers B (1996) Object-oriented metrics, measures of complexity. Prentice Hall

Hirohisa AMAN, Mochiduki N, Yamada H (2006) A model for detecting cost-prone classes based on Mahalanobis-Taguchi method. IEICE Trans Inf Syst 89(4):1347–1358

Hulse JV, Khoshgoftaar TM, Napolitano A, Wald R (2009) Feature selection with high-dimensional imbalanced data. In Proc. of International Conference on Data Mining Workshops, 507–514

Jeni L, Cohn JF, De La Torre F (2013) Facing imbalanced data–recommendations for the use of performance metrics. In Proc. of Humane Association Conf. on Affective Computing and Intelligent Interaction, 245–251

Kamei Y, Monden A, Matsumoto S, Kakimoto T, Matsumoto K (2007) The effects of over and under sampling on fault-prone module detection. In Proc. 1st International Symposium on Empirical Software Engineering and Measurement, 196–204

Khoshgoftaar TM, Seliya N, Sundaresh N (2006) An empirical study of predicting faults with case-based reasoning. Softw Qual J 14(2):85–111

Koru AG, Liu H (2007) Identifying and characterizing change-prone classes in two large-scale open-source products. J Syst Softw 80:63–73

Koru AG, Tian J (2005) Comparing high-change modules and modules with the highest measurement values in two large-scale open-source products. IEEE Trans Softw Eng 31(8):625–642

Kubat M, Matwin S (1997) Addressing the curse of imbalanced training sets: one sided selection. In Proc. of 14th International Conference on Machine Learning 97: 179–186

Lessmann S, Baesans B, Mues C, Pietsch S (2008) Benchmarking classification models for software defect prediction: a proposed framework and novel finding. IEEE Trans Softw Eng 34(4):485–496

Li M, Zhang H, Whu R, Zhou Z (2012) Sample-based software defect prediction with active and semi-supervised learning. Autom Softw Eng 19(2):201–230

Liu Y, An A, Huang X (2006) Boosting prediction accuracy on imbalanced datasets with SVM ensembles. In Advances in Knowledge Discovery and Data Mining, 107–118

Lopez V, Fernandez A, Garcia S, Palade V, Herrera F (2013) An insight into classification with imbalanced data: empirical results and current trends on using data intrinsic characteristics. Inf Sci 250:113–141

Lu H, Zhou Y, Xu B, Leung H, Chen L (2012) The ability of object-oriented metrics to predict change-proneness: a meta-analysis. Empir Softw Eng J 17(3):200–242

Malhotra R (2015) A systematic review of machine learning techniques for software fault prediction. Appl Soft Comput 27:504–518

Malhotra R, Khanna M (2013) Investigation of relationship between object-oriented metrics and change proneness. Int J Mach Learn Cybern. Springer-Verlag 4(4): 273–286

Malhotra R, Nagpal K, Upmanyu P, Pritam N (2014) Defect collection and reporting system for git based open source software. In Proc. of International Conf. on Data Mining and Intelligent Computing, 1–7

Martin RC (2002) Agile software development: principles, patters, and practices. Prentice Hall, USA

Menzies T, Greenwald J, Frank A (2007a) Data mining static code attributes to learn defect predictors. IEEE Trans Softw Eng 33(1):2–13

Menzies T, Dekhtyar A, Distefance J, Greenwald J (2007b) Problems with precision: a response to comments on 'data mining static code attributes to learn defect predictors'. IEEE Trans Softw Eng 33(9):637–640

Munkhdalai T, Namsrai OE, Ryu KH (2015) Self-training in significance space of support vectors for imbalanced biomedical event data. BMC Bioinf 16(7):1

Murphy KP (2006) Naive Bayes classifiers, Technical Report

Olague H, Etzkorn L, Gholston S, Quattlebaum S (2007) Empirical validation of three software metric suites to predict the fault-proneness of object-oriented classes developed using highly iterative or agile software development processes. IEEE Trans Softw Eng 33(10):402–419

Pai GJ, Dugan JB (2007) Empirical analysis of software fault content and fault proneness using Bayesian methods. IEEE Trans Softw Eng 33(10):675–686

Phua C, Alahakoon D, Lee V (2004) Minority report in fraud detection: classification of skewed data. SIGKDD Explorations 6(1): 50–59

Rodriguez D, Herraiz I, Harrison R, Dolado J, Riquelme JC (2014) Preliminary comparison of techniques for dealing with imbalance in software defect prediction. In Proc. of the 18th International Conf. on Evaluation and Assessment in Software Engineering, 43

Romano D, Pinzger M (2011) Using source code metrics to predict change-prone java interfaces. 27th IEEE International Conference on Software Maintenance, 303–312
Seiffert C, Khoshgoftaar TM, Hulse JV, Folleco A (2014) An empirical study of the classification performance of learners on imbalanced and noisy software quality data. Inf Sci 259:571–595
Seliya N, Khoshgoftaar TM (2011) The use of decision trees for cost-sensitive classification: an empirical study in software quality prediction. Wiley Interdiscip Rev: Data Min Knowl Disc 1:448–459
Shatnawi R (2012) Improving software fault-prediction for imbalanced data. In Proc. of International Conf. on Innovations in Information Technology, 54–59
Singh Y, Kaur A, Malhotra R (2009) Empirical validation of object-oriented metrics for predicting fault proneness models. Softw Qual J 18:3–35
Stone M (1974) Cross-validatory choice and assessment of statistical predictions. J R Soc A 36:111–114
Su CT, Hsiao YH (2007) An evaluation of the robustness of MTS for imbalanced data. IEEE Trans Knowl Data Eng 19(10):1321–1332
Tan M, Tan L, Dara S, Mayeux C (2015) Online defect prediction for imbalanced data. In Proc. of 37th International Conf. on Software Engineering
Visa S, Ralescu A (2005) Issues in mining imbalanced data sets- a review paper. In Proc. of 16th Conference on Artificial Intelligence and Cognitive Science, 67–73
Wang S, Yao X (2013) Using class imbalance learning for software defect prediction. IEEE Trans Reliab 62:434–443
Weiss GM (2004) Mining with rarity: a unifying framework. ACM SIGKDD Explor Newslett 6(1):7–19
Weiss GM, McCarthy K, Zabar B (2007) Cost-sensitive learning vs. sampling: which is best for handling unbalanced classes with unequal error costs?. In Proc. of International Conf. on Data Mining, 35–41
Weng CG, Poon J (2008) A new evaluation measure for imbalanced datasets. In Proc. of the 7th Australian Data Mining Conference, 27–32
Witten IH, Frank E, Hall MA (2011) Data mining: practical machine learning tools and techniques, 3rd edn. Morgan Kaufmann, San Francisco
Xu R, Chen T, Xia Y, Lu Q, Liu B, Wang X (2015) Word embedding composition for data imbalances in sentiment and emotion classification. Cogn Comput 7(2):226–240
Yang P, Yoo PD, Fernando J, Zhou BB, Zhang Z, Zomaya AY (2014) Sample subset optimization techniques for imbalanced and ensemble learning problems in bioinformatics applications. IEEE Trans Cybern 44(3):445–455
Zhang X, Li Y (2011) An empirical study of learning from imbalanced data. In Proc. of the 22nd Australasian Database Conf, 85–94
Zhou Y, Leung H, Xu B (2009) Examining the potentially confounding effect of class size on the associations between object metrics and change proneness. IEEE Trans Softw Eng 35(5):607–623



Dr. Ruchika Malhotra is Associate Head and Assistant Professor at the Department of Software Engineering, Delhi Technological University (formerly Delhi College of Engineering), Delhi, India. She has been awarded the prestigious UGC Raman Postdoctoral Fellowship by the Indian government for pursuing postdoctoral research from Department of Computer and Information Science, Indian University-Purdue University, Indianapolis, Indiana, USA. She was an assistant professor at the University School of Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India. She received her master's and doctorate degree in software engineering from the University School of Information Technology, Guru Gobind Singh Indraprastha University, Delhi, India. She has received IBM Faculty Award 2013. She has received Best Presenter Award in Workshop on Search Based Software Testing, ICSE, 2014, Hyderabad, India. Her h-index is 20 as reported by Google Scholar.

She is executive editor of Software Engineering: An International Journal. She is author of book titled "Empirical Research in Software Engineering" published by CRC Press, USA and co-author of a book on Object Oriented Software Engineering published by PHI Learning. Her research interests are in empirical research in software engineering, improving software quality, statistical and adaptive prediction models, software metrics, the search-based software engineering and software testing. She has published more than 130 research papers in international journals and conferences. She can be contacted by e-mail at ruchikamalhotra2004@yahoo.com.



Megha Khanna is currently pursuing her doctoral degree from Delhi Technological University. She is currently working in Sri Guru Gobind Singh College of Commerce, University of Delhi. She completed her master's degree in software engineering in 2010 from the University School of Information Technology, Guru Gobind Singh Indraprastha University, India. She received her graduation degree in computer science (Hons.) in 2007 from Acharya Narendra Dev College, University of Delhi. Her research interests are in software quality improvement, applications of machine learning techniques in change prediction, and the definition and validation of software metrics. She has various publications in international conferences and journals. She can be contacted by email at meghakhanna86@gmail.com.

Accepted Manuscript



An Energy Aware Cost Effective Scheduling Framework for Heterogeneous Cluster System

Neetesh Kumar, Deo Prakash Vidyarthi

PII: S0167-739X(17)30034-1
DOI: http://dx.doi.org/10.1016/j.future.2017.01.015
Reference: FUTURE 3297

To appear in: *Future Generation Computer Systems*

Received date: 21 June 2016
Revised date: 28 November 2016
Accepted date: 7 January 2017

Please cite this article as: N. Kumar, D.P. Vidyarthi, An Energy Aware Cost Effective Scheduling Framework for Heterogeneous Cluster System, *Future Generation Computer Systems* (2017), http://dx.doi.org/10.1016/j.future.2017.01.015

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

An Energy Aware Cost Effective Scheduling Framework for Heterogeneous Cluster System

Neetesh Kumar[1] and Deo Prakash Vidyarthi[2]

[1]Department of Computer Science and Engineering, Delhi Technological University, Delhi-110042, India

[2]School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi-110067, India

Abstract

Cloud computing is an emerging market place in which computing resources are treated as the utilities and priced for their usage. A huge competition prevails among the cloud Service Providers (SPs) to offer eco-friendly and economically viable cloud services. Heterogeneous Cluster System (HCS) forms the backbone of such computing resources and for eco-friendly services, minimal energy consumption in HCS is a challenging issue. It not only brings down the operational cost but enhance the system reliability and other environmental veneration as well. Commonly, a green Service Level Agreement (SLA) is made between cloud users and SP. Green SLA ensures that a certain percentage of $makespan$ of the schedule is sacrificed for an energy efficient schedule. This work proposes a scheduling framework to execute heterogeneous independent jobs on a HCS. It also introduces a SLA based negotiation scheme for makespan, energy and cost reffered as "Makespan, Energy and Cost Negotiator (MECN)". The proposed model utilizes all possible empty slacks of the schedule running on Processing Elements (PEs) and extended $makespan$ (due to SLA). Using Dynamic Voltage and Frequency Scaling (DVFS) technique, it provides an energy efficient schedule. A resource replacement strategy is also incorporated, in the proposed framework, to generate a cost effective schedule without sacrificing its performance. The performance of the model is analyzed by its simulation which reflects its usefulness and effectiveness.

Keywords: Heterogeneous Cluster System (HCS), Job Scheduling, Makespan, Green SLA, Dynamic Voltage and Frequency Scaling (DVFS)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1. Introduction

Over the past decade, a remarkable growth has been witnessed in cloud computing for its on demand IT services. The success of cloud computing is significantly due to Information and Communication Technology (ICT) and its capability to fulfill the worldwide demand for dedicated and customized services [1]. In offering these services the amount of energy consumed by the high-end computing systems, especially large-scale cluster system is significant. According to an estimate [1, 2], high-end computing facilities in the cloud systems consume 2% of world's electricity which is likely to grow four fold by 2020. High energy consuming resources e.g. datacenters, generate more heat requiring cooling devices which eventually emit intolerable carbon harming the environment [4]. This raises very serious concerns regarding economic viability, environment and resource sustainability in the near future [5, 6]. Therefore, scientists emphasize for better energy consumption of these High Performance Computing (HPC) data centers and consider it an important research topic [3].

Modern computing devices (processors/cores/PEs) are well-equipped with the DVFS technology [5, 7] which allows the computing resources to operate at multiple levels of voltage and frequency. DVFS technique is applied for minimal energy consumption by scaling down the voltage and the frequency levels of the involved PEs. This slows down the computation speed of the PEs, affecting the $makespan$ of the schedule. To overcome this, a green SLA can be made between a client and SP which allows the client to compromise the performance ($makespan$) a bit in lieu of an energy efficient schedule. Moreover, it has been observed that it is nontrivial to schedule the heterogeneous jobs to the heterogeneous PEs uniformly. In other words, some empty slacks remain unused on most of the PEs due to heterogeneous nature of jobs and PEs [3, 20]. A possibility arises to minimize the energy consumption on HCSs using these slacks. For this, DVFS technique can be applied to scale down the supply voltage and frequency levels of the PEs for the available slack periods.

A good number of models have been proposed by the researchers that utilizes the DVFS technique to reduce the energy consumption of the HPC systems [11-15]. In general, to offer an energy efficient schedule at a lower cost is strategically desired by the SP to attract more number of customers. Most of the recent works, in the domain of energy and cost saving, consider variable energy price, peak energy charge and energy consumed by the cooling systems to reduce electricity cost for the HPC cloud providers. The fundamental idea, of these energy minimization models, is to utilize the slack time in order to scale down the supply voltage and operating frequency of the computing resources [3, 11-15, 20]. Further, most of the research to minimize energy prices is separately done by considering various factors e.g. peak energy charges, energy consumed by the cooling system etc. [1].

ACCEPTED MANUSCRIPT

Unlike others, the proposed work utilizes the heterogeneity characteristics of the available PEs to provide an energy efficient as well as cost effective schedule. As obvious, PEs operating on higher voltage and frequency levels (provide fast computation) are offered at higher prices ($ per time unit). On the contrary, PEs operating on lower voltage and frequency levels (provide slow computation) are offered at low prices. In general, client initially demands PEs operating at the high voltage and frequency for the faster computation. Later, with green SLA, client has to sacrifice some performance, and faster PEs slow down by scaling down their voltage and frequency levels using DVFS technique. Also, it is observed that a client has to pay more for the high speed PEs which are running on low speed (poorly utilized). An opportunity exits to replace the high speed and priced ($) PEs with other available low speed and priced ($) PEs in the cluster. In order to replace PEs, it is ensured that PEs operate at appropriate voltage and frequency levels while satisfying the SLA and offer comparatively low prices.

The objective, in this work, is to offer a combined Energy Aware and Cost Effective Framework (EACEF) in presence of SLA without compromising the Quality of Service (QoS). In this framework, clients negotiates with the SP for the QoS by requesting for minimal energy consumption on a given $makespan$ extension rate. Further, EACEF utilizes the heterogeneity characteristics of PEs and jobs to explore the available slacks [3]. During slack periods, DVFS technique is used to minimize the energy consumption. A resource replacement module, an important component of the model, offers an economical schedule as per resource availability. A simulation study of EACEF is conducted on a heterogeneous cluster using three different real PEs [20]. A workload generator is used to generate heterogeneous workload of some common mathematical and computational problems [8, 37]. The performance of the proposed model is evaluated and is compared with two well-known heterogeneous schedulers i.e., best effort scheduler (HEFT) [9] and ESHMP [10].

The significant contributions, of the proposed model, are listed as bellow.

- It is a green SLA based novel model which minimizes energy and cost to generate an economic and effective schedule for heterogeneous cluster system. It utilizes the heterogeneity characteristics of both (PEs and jobs) to minimize energy consumption.
- It also utilizes empty slacks of the PEs created due to heterogeneity characteristics of the PEs, jobs and other slacks created due to green SLA.
- A resource replacement mechanism is also introduced to replace the higher costing PEs with lower cost PEs (as per availability). This offers a cost effective schedule without compromising the QoS.

ACCEPTED MANUSCRIPT

- A simulation study is done to study the effectiveness of EACEF over two other heterogeneous schedulers i.e., HEFT and ESHMP. The comparative results show its effectiveness.

Rest of this paper is organized as follows. In section 2, some related work is presented. The system model is introduced in section 3 whereas problem formulation is given in section 4. In section 5, the proposed model is discussed. Simulation experiments for the comparative analysis is done in section 6. Finally, the work is concluded in section 7.

2. Related Work

DVFS technique, along with SLA, has been widely used in cluster/cloud computing environment to minimize the energy consumption [11-18]. SLA is an agreement between the SP and the clients to meet various QoS parameters such as computational cost, deadline, response time and other distributed services [12, 13, 17, 18]. Wu et al. [11] proposed a SLA based technique for computing resource allocation. In this, a set of jobs is assigned to fewer DVFS enabled processing units under SLA. For doing so, frequencies of the processing units are scaled down to the appropriate level for the execution. Gao et al. [13] also proposed a SLA based approach for dynamic resource management. This satisfied response time by utilizing both; server consolidation and DVFS to minimize energy consumption. Costa et al. [17] proposed a multi-facet technique to optimize the energy consumption in grids and clouds. They considered user's decisions and SLA-aware approach for cloud resource management [18].

A literature survey done in [1] depicts that very meagre research work focuses on energy and cost minimization. In [21], authors considered variable energy prices, peak energy charge, and energy consumed by the cooling systems for reducing the electricity charges. Their work was dedicated for HPC Cloud with multiple geographically distributed Data Centers (DCs). Bila et al. [22] described partial VM migration for desktops, motivating its usefulness as an energy saving technique. In particular, they observed that their technique can reduce annual energy costs in overnight hours for the companies with 100 or more desktops by atleast 55%. They referred to a third-party Infrastructure as a Service (IaaS) cloud, an intranet backend, or even a federation of other PCs in the same domain, but not to a cloud federation. Further, Goiri et al. [23] presented Parasol (a prototype green DC including a set of solar panels) and green switch (an approach to dynamically schedule the workload and to select the energy sources to use) in order to demonstrate that intelligent workload and energy source management can reduce energy costs. Hongxing Li et al. [24] also introduced an efficient model for the resources in a federation of geo-distributed clouds. In particular, authors optimized the delay for each job in order to maximize the profit for the cloud. However, they did not focus on how to minimize the energy costs. In

ACCEPTED MANUSCRIPT

[25], the authors have presented a very interesting policy to address energy and cost saving in federated Cloud. They considered three cost factors: energy price, peak energy charge, and the energy consumed by the cooling system and concluded that the cost aware load placement policy can significantly push down provider's operational cost.

Buyya et al. [26] asserted that the next generation cloud SPs should be able to dynamically expand or resize their provisioning capability. They evoked the factor that poses difficult problems in effective provisioning and delivery of application services. They suggested the application of market-based utility to reduce energy usage. In [27], Koch et al. introduced a method based on Maximum Likelihood Estimation (MLE) considering the heterogeneity of infrastructure and the resource allocation plans. This maximizes the platform utilization for education and cost reduction as compared with conservative resource allocation strategies. Paul et al. [28] suggested energy aware load distribution and electricity cost volatility control for cloud SPs. They investigated load distribution strategies to minimize electricity cost in compliance with service level agreement (SLA). It also considers the adverse effects of switching the servers. Buyya et al. [29] proposed an energy efficient cloud service pricing: a two-timescale optimization approach. They provided an analytical framework by considering the operational cost incurred by the CSP to service the given demand.

Conspicuous is the uses of DVFS technique for minimizing the energy consumption of the computing resources. In using DVFS, in general, there is energy and makespan tradeoff. Most of the energy minimization models talk about the electricity cost via different approaches like energy price, peak energy charge, and energy consumed by the cooling systems etc. This work introduces a different multi-objective scheduling framework considering energy and performance tradeoff in heterogeneous cluster systems. It utilizes DVFS technique along with green SLA to minimize the energy consumption and through scheduling, it also offers a low priced computation without compromising its performance.

3. The System Model

The proposed EACEF model is comprised of a green SLA based HCS architectural model, application/job structure model and energy model discussed as follow.

3.1 Green SLA based HCS Architectural Model

The outline of the proposed EACEF model in a heterogeneous cluster environment is depicted in Fig. 1. In general, the green SLA-based system model is classified in three layers: client layer, green SLA based scheduler layer, and SP layer. A client submits his/her heterogeneous jobs (heterogeneity in terms

of job type and size). The scheduler layer generates energy efficient and cost effective schedule for the jobs, along with its mapping on the cluster nodes while satisfying negotiation conditions. The objective of the proposed model is to convert a schedule, obtained from a base scheduler, into an energy efficient and cost effective schedule in the presence of a green SLA.

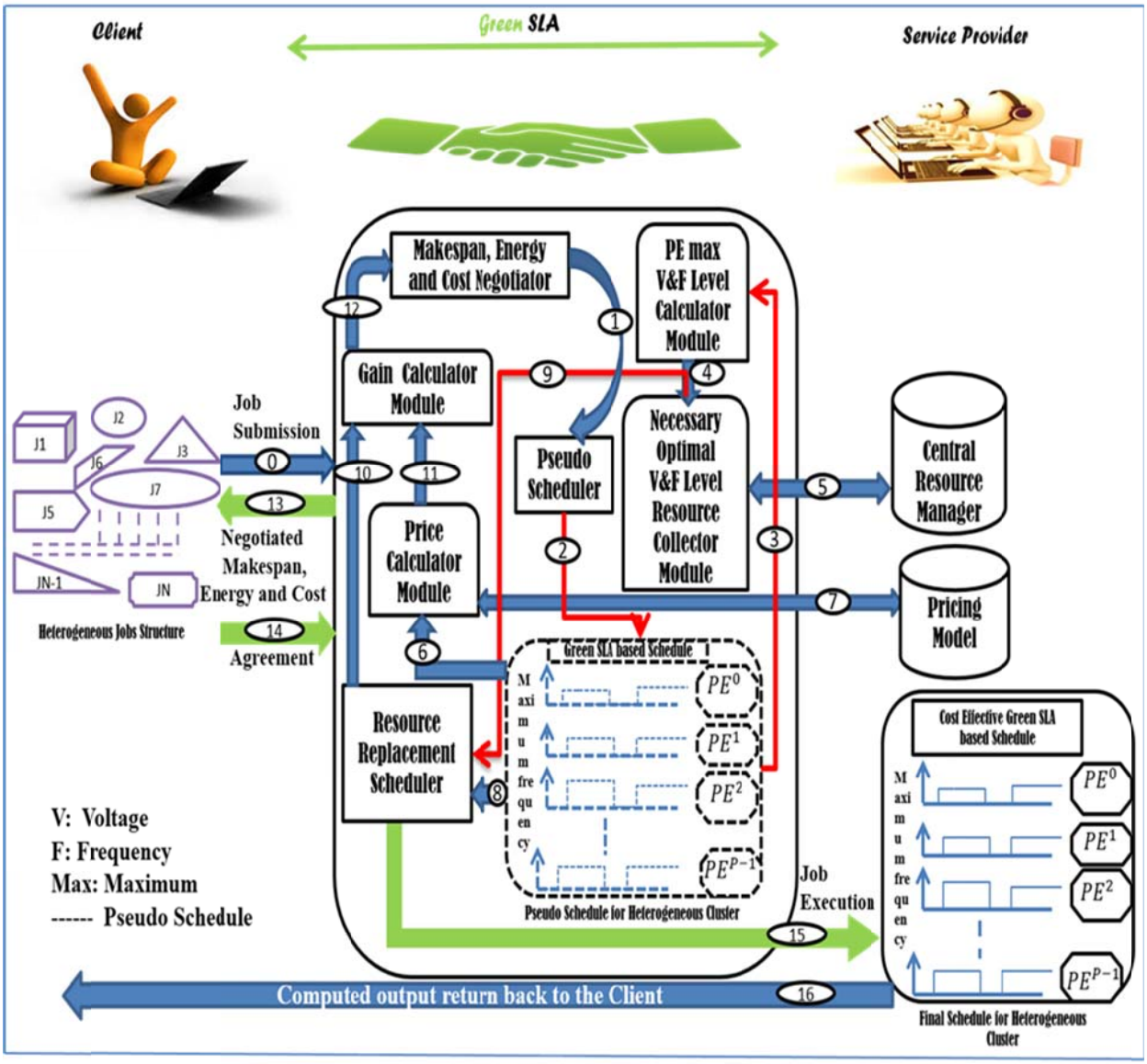


Fig. 1 A green SLA-based EACEF for a HCS

To perform this, all the steps are numbered [0 to 16] in Fig. 1. In step 0, a client submits his/her jobs to the scheduler layer. At the scheduler layer, a base scheduler (i.e., HEFT/ESHMP) maps these jobs to the PEs and generates a $Base_Schedule$. $Base_Schedule$ is suppose to generate better $makespan$.

ACCEPTED MANUSCRIPT

Next, *Base_Schedule* is submitted to the "Makespan, Energy and Cost Negotiator (MECN)" to negotiate makespan, energy and cost. Applying an ordered subsequent step numbers [1-12] on the *Base_Schedule*, makespan, energy and cost are negotiated at the scheduler layer. The MECN transfers *Base_Schedule* to the pseudo scheduler, through step. As an output of pseudo scheduler, a *Pseudo_Schedule* (energy aware) is generated in step number 2. In step 3, this *Pseudo_Schedule* becomes an input to the PE "Maximum Voltage & Frequency Calculator Module (MVFCM)" is used to calculate maximum voltage and frequency of the used PE. The required maximum voltage and frequency (observed in step 3) is forwarded to "Necessary Optimal Voltage & Frequency Level Resource Collector Module (NOVFLRCM)" as an input in step 4. In step 5, NOVFLRCM collects the minimum fulfilling required voltage and frequency available resources from "Central Resource Manager (CRM)".

In steps 6 and 7, *Pseudo_Schedule* and list of pricing rates (from pricing model) are given as inputs to the "Price Calculator Module (PCM)" to calculate the cost of the *Pseudo_Schedule*. In steps 8 and 9, *Pseudo_Schedule* and available lower voltage and frequency operating PEs (collected by NOVFLRCM) are given as inputs to the "Resource Replacement Scheduler (RRS)". RRS module replaces the higher cost PEs with lower cost PEs and generates final schedule i.e., "*Final_Schedule*". The information of the replaced PEs in *Final_Schedule* and cost of pseudo schedule (calculated by PCM) with PEs rate list are transferred to the "Gain Calculator Module (GCM)" in steps 10 and 11 respectively. Now, GCM calculates the gain and sends it to the MECN module using step 12. Through step 13, the negotiated makespan, energy and cost is returned back to the client. After a self decision, client confirms their agreement to the scheduler layer in step 14. With the same agreement, scheduler layer sends the scheduled jobs (*Final_Schedule*) to the SP layer in step 15. SP layer computes the schedule using the same resources as suggested by the scheduler layer. After succesfull execution, results are directly communicated back to the client in step 16. Various modules, used in EACEF, are briefed as follows.

MECN: Before a client uses services, he/she first negotiates with SP for some performance metrics i.e., makespan, energy and cost. The information of the jobs and PEs are forwarded to the scheduler layer. A module, named "Makespan, Energy and Cost Negotiator", negotiates on the *makespan* and energy consumption tradeoff. With this tradeoff, this module also offers a cost effective (low price) schedule as per the availability of the resources (PEs). After a successful negotiation, an agreement is reached in the form of % makespan extension and % dropped computational cost (price). For example, an agreement can be made such as an increase in 15% of job execution time is allowed to reduce some energy consumption. Also, the client is charged 5% lesser price.

ACCEPTED MANUSCRIPT

PS: The PS is responsible for transforming a $Base_Schedule$ (generated by HEFT/ESHMP etc.) into an energy efficient schedule ($Pseudo_Schedule$) using DVFS technique followed by green SLA.

MVFC: This module calculates the maximum voltage and frequency levels used by respective PEs for the $Pseudo_Schedule$ produced by the pseudo scheduler.

NOVFLRC: This module communicates to the Central Resource Manager (CRM) and collects those available PEs in the heterogeneous cluster; that are having their maximum voltage/frequency levels and are least suitable to execute the $Pseudo_Schedule$.

PC: This module estimates the total cost of the PEs used by $Pseudo_Schedule$. Pricing rates of the PEs are collected from pricing model as accessible at the SP's side.

RRS: RRS replaces resources (PEs) of the $Pseudo_Schedule$ as per their availability as confirmed by NOVFLRCM and produces a $Final_Schedule$.

GC: GC is used to compute the % gain in the cost ($) of $Final_Schedule$ over the $Pseudo_Schedule$.

Finally, SP layer consisting of various heterogeneous PEs, is responsible for job execution following the $Final_Schedule$ as recommended by the scheduler layer. The required PEs are explicitly provided to the user for a scheduled duration ($makespan$) on best (lower) price as suggested by EACEF. After successful execution, results are reported back to the client.

3.2 The Heterogenous Cluster Model

This subsection formulates the HCS comprised of multiple heterogeneous Processing Elements (PEs) interconnected via communication links, as shown in Fig. 1. It is expected that each PE in the cluster is DVFS enabled. A DVFS enabled PE can operate on multiple levels of supply voltage (V) and operating frequency (F). That is formally represented by Eqs. 1 and 2.

$$V = \bigcup_{1 \le l \le M} \{v_l\} \tag{1}$$

$$F = \bigcup_{1 \le l \le M} \{f_l\} \tag{2}$$

Where, M denotes the level of maximum operating voltage and frequency of the respective PE. Further, it is assumed that number of PEs in a cluster are P (PE_{set}) which can be represented as a set of PEs using Eq. 3.

ACCEPTED MANUSCRIPT

$$PE_{set} = \bigcup_{0 \le j \le P-1} \{PE^j\} \tag{3}$$

The heterogeneity characteristic of the cluster can be represented by Eq. 4 which denotes that HCS is truly heterogeneous. Two or more PEs, with same voltage and frequency levels, are allowed in the cluster system and the proposed model works without any deterioration in the performance.

$$\begin{cases} \left(\left(f_l^j \neq f_l^{j'} \right) \middle| \left(f_l^j = f_l^{j'} \right) \right) \\ \left(\left(v_l^j \neq v_l^{j'} \right) \middle| \left(v_l^j = v_l^{j'} \right) \right) \end{cases} \quad where, j \neq j' and\ 0 \le j, j' < P,$$

$$1 \le l \le m \ and\ v_l^j,\ v_l^{j'} \in V \ \ and\ f_l^j, f_l^{j'} \in F \tag{4}$$

Where, v_l^j and f_l^j denote l^{th} level of voltage and frequency respectively on j^{th} PE.

3.3 The Application Model

In this work, an application is considered to consists of N number of heterogenous jobs (job structure with different types and sizes of symbols as shown in Fig. 1) $i.e. J_{set} = \{J_1, J_2 \ldots . J_N\}$. Since, jobs are submitted by the client to execute on a cluster system comprising of P numbers of hetrogenous PEs (as discussed in section 2.2) represented as $PE_{set} = \{PE^0 PE^1 \ldots \ldots PE^{P-1}\}$. In general, $N \geq P$. The nature of jobs is described by Eqs. 5 and 6.

$$\begin{cases} \{size(J_i) = size(J_{i'})\} | \\ \{size(J_i) \neq size(J_{i'})\} \end{cases} \tag{5}$$

$$\begin{cases} \{type(J_i) = type(J_{i'})\} | \\ \{type(J_i) \neq type(J_{i'})\} \end{cases} \tag{6}$$

Eqs. 5 and 6 mathematically depict the nature of the jobs. Where, $J_i, J_{i'} \in J_{set}$, and $i \neq i'$. The $size(J_i), type(J_i)$ denote the size of job J_i (number of instructions in J_i) and type of application for job J_i, respectively. Eq. 5 describes that two or more jobs can be of same size or different size. Similarly, Eq. 6 describes that two or more jobs can be from same or different applications. In summary, the jobs are not constrained by their size and type. Eqs. 5 and 6 can have the inference as in Eq. 7.

$$\begin{cases} \{ {t_{exec}}_i^{\ j} = {t_{exec}}_i^{\ j'} \} | \\ \{ {t_{exec}}_i^{\ j} \neq {t_{exec}}_i^{\ j'} \} \end{cases} \tag{7}$$

ACCEPTED MANUSCRIPT

Where $t_{exec_i}^{j}$ denotes the execution time of i^{th} job on j^{th} PE ($j \neq j'$) and is calculated using Eq. 8. Eq. 7 states that the execution time, on two PEs of the cluster system, can be same or different. It depends on the characterstics of the jobs and PEs.

$$t_{exec_i}^{j} = \frac{w(t_i) \times CPI^j}{f_{max}^j} \tag{8}$$

Where, CPI^j is the number of cycles per instruction on j^{th} PE. It can be determined by both the hardware and software [3]. f_{max}^j denotes the frequency of j^{th} PE running at its maximum frequency.

3.3 The Energy Model

The total energy consumption (EN) of modern DVFS enabled PE is defined as the sum of Dynamic Energy Consumption($EN_{dynamic}$), and Static Energy Consumption (EN_{static}) [31] as shown in Eq. 9. In general, dynamic energy consumption is due to charging and discharging of the circuit node capacitance in the logic gates. Static energy consumption is due to bias and leakage of current [3].

$$Total\ Energy\ Consumption\ (EN) = EN_{dynamic} + EN_{static} \tag{9}$$

Dynamic Energy Consumption: It is presumed that all the PEs in the cluster are designed using CMOS technology. Thus, dynamic energy consumption ($Power_{dynamic}$) of CMOS circuits is calculated using Eq. 10.

$$Power_{dynamic} = N_{sw} C_L {V_{dd}}^2 f = C_{eff} {V_{dd}}^2 f \tag{10}$$

Where, $C_{eff} = N_{sw} C_L$, N_{sw} and C_L denote average switched capacitance per cycle, average gate switching factor at each clock cycle and loading capacitance of crystal respectively. Supply voltage (V_{dd}) and frequency (f) relationship is also shown in Eq. 11 [19].

$$f = \frac{(V_{dd} - V_{th})^{\alpha}}{K L_d} \tag{11}$$

From Eq. 11, it is obvious that frequency increases in proportion to the supply voltage. The supply voltage can be calculated using Eq. 12 if the operating frequency is known.

$$V_{dd} = V_{th} + (K L_d F)^{\frac{1}{\alpha}} \tag{12}$$

ACCEPTED MANUSCRIPT

Here, K, L_d and α are constant values of which depend on the technology used in designing the PE. Table 1 represents the constant parameters for 0.18 μm technology. The constant values are derived assuming that the voltage and frequency are specified in volts and hertz respectively.

Briefing Eqs. 10, 11 and 12, $EN_{dynamic}$ is significantly affected by the supply voltage. Therefore, the reduction in voltage is useful to lower the energy consumption. Hence, for N number of jobs using P number of PEs, the $EN_{dynamic}$ can be defined as in Eq. 13.

Table 1 Values for 0.18 μm technology [19]

Constant	Value
C_{eff}	1.11*10^{-9} F
K	51*10^{-12}
L_d	37
α	1.5

$$EC_{dynamic} = \sum_{1\le i\le N,0\le j\le P-1,1\le l\le M} PC_{dynamic} \times t_{exec}{}^{j}_{i,l} = \sum_{1\le i\le N,0\le j\le P-1,1\le l\le M} C_{eff}(v_l^j)^2 f_l^j \times t_{exec}{}^{j}_{i,l} \quad (13)$$

In Eq. 13, $t_{exec}{}^{j}_{i,l}$ denotes the execution time of a task t_i on j^{th} PE which is running on its l^{th} level of voltage and frequency pair. Where, $1 \le l \le M$ and $v_l^j \in V$, $f_l^j \in F$. P denotes the maximum number of PEs in the cluster and M denotes the maximum number of voltage and frequency levels of the respective PE in the cluster.

Static Energy Consumption: It is known that PEs consume some amount of energy, even in an idle state, called static energy consumption (EN_{static}). The static power ($Power_{static}$) consumption of the PE can be calculated using Eq. 14.

$$Power_{static} = C_{eff} {V_{th}}^2 f_{min} \quad (14)$$

Thus, EN_{static} for P number of PEs in a cluster can be derived by Eq. 15.

$$EN_{static} = \left(\sum_{0\le j\le P-1}\left(C_{eff}(v_1^j)^2 f_1^j\right) \times mkespan\right) - \sum_{1\le i\le N,0\le j\le P-1,}\left(C_{eff}(v_1^j)^2 f_1^j \times t_{exec}{}^{j}_{i,1}\right) \quad (15)$$

In Eq. 15, $t_{exec}{}^{j}_{i,1}$ denotes the execution time of task t_i running on j^{th} PE which is operating on its $first$ level of voltage and frequency.

ACCEPTED MANUSCRIPT

4. The Motivation and the Problem Formulation

A heterogeneous cluster has a number of heterogeneous PEs along with their distinct operating voltage and frequency levels. Heterogeneous PEs are offered to the client at different pricing rates [32]. A high speed (operating on higher voltage and frequency) PE has higher cost ($). It is very logical to note that when DVFS technique is applied for a schedule to minimize energy consumption, it slows down the operating voltage and frequency levels of PEs. When high speed PEs are not being utilized properly, despite of high voltage and frequency levels, PEs run on their respective low voltage and frequency levels. This makes their operating cost per time unit very high. An opportunity exists to replace these high cost PEs with the low cost PEs operating at appropriate operating voltage and frequency. By doing so, a service can be delivered to the client at lesser cost without compromising the performance. In brief, it is not worth to use high-speed PEs on higher rates for the computation if the PEs are not being utilized at their maximum voltage and frequency. Those PEs can be used for other highly demanded faster computations. Instead, it is better to use less costly PEs for the computation with appropriate voltage and frequency levels. This observation motivates us to design a model that generates an economical (cost effective) schedule considering $makespan$ and energy performance tradeoff under the influence of green SLA.

In this work, a problem of energy-performance tradeoff with green SLA is considered which provides a cost effective schedule for a HCS. The problem is to minimize the energy consumption and cost ($) of a schedule (mapping of *N* number of heterogeneous jobs of an application to *P* number of PEs) within the limitation of a green SLA. A mathematical formulation, for the above stated problem, is designed using Eqs. 16, 17, 18 and 19.

$To\ minimize\ ,$

$$Total\ Energy\ Consumption\ (EC) = EC_{dynamic} + EC_{static} = \left(\sum_{1\le i\le N, 0\le j\le P-1, 1\le l\le M} C_{eff}(v_l^j)^2 f_l^j \times t_{exec}{}_{i,l}^{j}\right) + \left(\sum_{0\le j\le P-1}\left(C_{eff}(v_1^j)^2 f_1^j\right) \times mkespan\right) - \sum_{1\le i\le N, 0\le j\le P-1}\left(C_{eff}(v_1^j)^2 f_1^j \times t_{exec}{}_{i,1}^{j}\right) \quad (16)$$

$Subject\ to$

$$makespan <= makespan \times (1+\mu) \quad (17)$$

$$Cost_{Final_Schedule} <= Cost_{Pseduo_Schedule} \quad (18)$$

$$Cost_{Schedule} = \sum_{j=0}^{P-1} makespan \times \$_per_time_unit(PE^j) \quad (19)$$

ACCEPTED MANUSCRIPT

The constraints of the time and cost are specified by Eqs. 17 and 18 respectively. The cost of the Final_Schedule and *Pseudo_Schedule* can be estimated using Eq. 19 with the $*_per_time_unit* price rate of used respective PE.

5. The Proposed Model

This section presents the detailed description of the proposed framework (EACEF) along with its various modules, shown in Fig.1.

The Proposed Framework (EACEF**)**

Input:

J_{set}: $\{J_1, J_2 \dots J_N\}$; // *N numbers of jobs*;

PE_{set}: $\{PE^0 PE^1 \dots PE^{P-1}\}$; // *P number of PEs*;

Base_Schedule: The output schedule of the base scheduler ; // *Base scheduler* (*i.e.*, *HEFT*, *ESHMP*)

μ: Maximum acceptable makespan extension; // *taken in form of* % (*i.e.*, 10%, 15% *...etc.*);

F_{max}: Maximum frequency and voltage pairs of each PE in P;

Output:

Computed result to the client;

Begin

Initialize: Flag = False;

[Final_Schedule, Makespan, Saved_Energy, Gain] = MECN(J_{set}, PE_{set}, Schedule, μ, F_{max});

Send back Makespan, Saved_Energy and Gain to the client for the negotiation;

Flag = Client_Agreement (); //*Client confirms negotiation agreement via Client_Agreement* ();

If (Client_Agreement == True)

Forward Final_Schedule to the SP for the computation;

After computation, release the resources of SP and directly returns the" computed result to the client";

End If

End

EACEF (inside of green SLA based scheduler layer) begins with number of inputs i.e., J_{set}, PE_{set}, *Base_Schedule*, μ and F_{max}. Initially, one boolean variable i.e., *Flag* is taken to capture the client response which is initialized by "False" value. Next, *MECN* module of the proposed EACEF is called. In general, *MECN* is responsible to establish the dataflow link among the other modules of EACEF, and results in the form of *Final_Schedule, Makespan, Saved_Energy and Gain*. The *Final_Schedule* is the energy efficient, cost effective schedule, and *Makespan, Saved_Energy and Gain* are the best offering values by EACEF which are yet to be agreed by the client for the further computation. Thus, *Makespan, Saved_Energy and Gain* are communicated to the client for the negotiation. The client response in boolean value is captured in *Flag* variable by

$Client_Agreement()$ which is used to denote the client response (True/False) only. If, it is $True$ then $Final_Schedule$ is forwarded to the Service Provider (SP) for the computation as per schedule. After successful execution, used resources are released and output (result) is returned back to the client.

5.1 Makespan, Energy and Cost Negotiator (MECN)

As mentioned, it is obvious to sacrifice some $makespan$ for energy etc. This module is used to negotiate the $makespan$ extension rate (μ) between client and SP to maintain an energy and $makespan$ tradeoff via green SLA. With this, it also offers a lower price computation as per the availability of appropriate resources (PEs), during the negotiation.

Module 1: MECN

Input:

J_{set}, PE_{set}, Base_Schedule, μ, F_{max};

Output:

Final_Schedule; // *Energy efficient and cost effective schedule (from RRS)*;

Makespan // *Final_Schedule updated makes pan*;

Saved_Energy; // *Calculated saved energy of the Pseudo_Schedule using Energy Model*;

Gain //*Gain in cost of Final_Schedule over Pseudo_Schedule*;

Begin

Calculate: Makespan $= \max_{1\le i\le N;0\le j<P} comp_{time_{i,j}}$;

//comp_time$_{i,j}$ denotes computation time of i^{th} job on $j^{th}PE$;

Calculate: Mak_extn $=$ Makespan $\times \frac{\text{Mak_extn}}{100}$;

[Pseudo_Schedule, Frqcy_matrix, Saved_Energy] $=$ PS(J_{set}, PE_{set}, Schedule, Makespan, Mak_extn, μ, F_{max});

[PE_Max_VF,] $=$ MVFC(J_{set}, PE_{set}, Pseudo_Schedule, Frqcy_matrix);

[New_PE_Set] $=$ NOVFLRC(PE_{set}, F_{max}, PE_Max_VF, Avail_PE);

[Pseudo_Total_Cost] $=$ PC(PE_{set}, makespan, Rate_List);

[Final_Schedule, Final_New_PE_Set] $=$ RRS(PE_{set}, Pseudo_Schedule, New_PE_Set);

[Gain] $=$ GC(Makespan, Final_New_PE_Set, Pseudo_Total_Cost, Rate_List);

End

It begins by taking a $Base_Schedule$ as one of the input (obtained from base scheduler) along with other inputs as listed in module 1. First, it calculates the $makespan$ and Max_extn for the $Base_Schedule$. By doing so, it makes call for different modules of EACEF along with their input and output parameter values as shown in modules. During the module/procedure call, left side (in […]) and right side (in (…)) parameters show the outputs and inputs respectively. Through the subsequent

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

procedures call (as shown in module 1), the MECN produces $Final_Schedule$, $Makespan$, $Saved_Energy$ and $Gain$ as the outputs: $Makespan$, $Saved_Energy$ and $Gain$ are used for the negotiation and $Final_Schedule$ is used for the computation only after, an agreement is reached between client and SP.

5.2 Pseudo Scheduler (PS)

Next module is to minimize the energy consumption of the $Base_Schedule$ using DVFS technique and considering green SLA. This module works on the PEs which are dedicatedly assigned by the base scheduler (HEFT/ESHMP). Note that, the used PEs in the $Base_Schedule$ are best PEs (served at a higher cost) to minimize $makespan$ of the schedule. Once, the PS module is applied on the $Base_Schedule$, these PEs may not be utilized properly. The fact that DVFS technique scales down the voltage and frequency levels of the PEs, gives an opportunity to replace these high cost PEs with the lower cost PEs (as per availability) to generate an energy efficient and cost effective $Final_Schedule$. The required list of inputs for PS is given in module *2*. This module begins by calculating the Actual Start Time (AST_i^j) and Actual Finish Time (AFT_i^j) [20] of i^{th} job which is running on respective j^{th} PE. Where, $1 \leq i \leq N$ and $0 \leq j < P$.

Next, for each j^{th} PE, the empty slacks for the $Base_Schedule$ are calculated ($i.e., Slack_PE^j$) without considering Mak_extn. This is formally a difference between $makespan$ and AFT of last job running on j^{th} PE. Further, it is easy to calculate the total slack of the j^{th} PE ($Total_Slack_PE^j$) by adding Mak_extn and $Slack_PE^j$. It is possible to calculate new operating frequency of j^{th} PE ($New_Freq_PE^j$) using a formula as shown in module 2. Now, for the $New_Freq_PE^j$, a best-fit level of voltage and frequency of the respective j^{th} PE is estimated. Further, new execution time ($t'^j_{exec_i}$) of each job running on the same PE is calculated. With $t'^j_{exec_i}$, new updated Actual Start Time (AST'^j_i) and Actual Finish Time (AFT'^j_i) are also calculated for the respective job. Using AST'^j_i and AFT'^j_i, the slots of the $Base_Schedule$ are updated from $[AST_i^j, AFT_i^j]$ *to* $[AST'^j_i, AFT'^j_i]$.

With this, current operating voltage and frequency level of i^{th} job running on j^{th} PE are recorded in a matrix $Frqcy_matrix(PE_i^j)$. At the end, an energy efficient $Base_Schedule$ is obtained which is transformed into $Pseudo_Schedule$. Using the energy model (section 3.3), total energy saving of $Pseudo_Schedule$ over $Base_Schedule$ is overserved and stored in Saved_Energy. This way, *PS* ensures its outputs $Pseudo_Schedule$, $Frqcy_matrix(\dots)$ and Saved_Energy. These outputs are used in succeeding modules to generate $Final_Schedule$.

ACCEPTED MANUSCRIPT

Module 2: PS

Input:

J_{set}, PE_{set}, Base_Schedule, Makespan, Mak_{extn}, μ, F_{max};

Output:

Pseudo_Schedule; // An energy efficient job schedule;

Frqcy_matrix(…);

//Stores updated voltage and frequency level of i^{th} job running on j^{th} PE in Pseudo_Schedule, where, $1 \le i \le N, 0 \le j \le P-1$;

Saved_Energy; *// Calculated saved energy of the Pseudo_Schedule using Energy Model*;

Begin

For (i = 1 to N) **do**

Calculate: AST_i^j and AFT_i^j from the Base_Schedule given by HEFT;

//AST and AFT of i^{th} job on its scheduled j^{th} PE;

End for

For (j = 1 to P − 1) **do**

Calculate: $Slack_PE^j = makespan - AFT_{end}^j$;

Calculate: $Total_Slack_PE^j = Slack_PE^j + Mak_extn$;

Calculate: $New_Freq_PE^j = \left(\frac{AFT_{end}^j - AST_1^j}{(AFT_{end}^j - AST_1^j) + Total_Slack_PE^j} \times fmax^j \right)$; // $fmax^j \in F_{max}$;

Map: The $New_Freq_PE^j$ to the best fit voltage and frequency level of the j^{th} PE;

for (i = 1 to length(PE^j)) **do** *// $length(PE^j)$ denotes number of jobs allocated on $j^{th}PE$*;

Calculate: the $t'^j_{exec_i} = \frac{w(t_i) \times CPI^j}{New_Freq_PE^j}$;

Calculate: $AST'^j_i = AFT^j_{i-1}$ and $AFT'^j_i = AST'^j_i + t'^j_{exec_i}$;

Reschedule the time slot of Base_Schedule for i^{th} job from $[AST_i^j, AFT_i^j]$ to $[AST'^j_i, AFT'^j_i]$;

Update the current operating voltage and frequency level of i^{th} job running on j^{th} PE of Base_Schedule and store in Frqcy_matrix(PE_i^j);

End for

End for

Makespan = makespan + Mak_extn; *// makespan is updated*;

Pseudo_Schedule = Base_Schedule; *// an updated Base_Schedule as Pseudo_Schedule*;

Saved_Energy = calculate saved energy of Pseudo_Schedule using Energy Model (section 3.3);

End

ACCEPTED MANUSCRIPT

5.3 Maximum Voltage Frequency Calculator (MVFC)

This module calculates and collects the maximum current operating voltage and frequency level of each PE, scheduled to execute the jobs. The forthcoming modules uses this information to find best fit available PEs at the SP site. This is desired, for resource replacement purpose, to generate a cost-effective schedule. Initially, this module creates P number of job sets (one for each PE) by separating the jobs running on respective PEs. Next, it selects the highest voltage and frequency levels of the jobs running on maximum operating voltage and frequency level of j^{th} PE in the j^{th} job set ($Jset(j)$) and stores in $PE_Max_VF(j)$. Where, $0 \leq j < P$.

Module 3: MVFC

Input:

J_{set}, PE_{set}, Pseudo_Schedule, Frqcy_matrix ;

Output:

PE_Max_VF: Each PE's maximum operating voltage and frequency levels after applying DVFS technique

Begin

For (i = 1 to N) **do**

$Jset(j) = Jset^{j} \cup Job_{i}^{j}$;

// Create job sets for each j^{th} PE, where, i^{th} job is running on j^{th} PE at respective voltage and frequency level of the PE in Pseudo_Schedule;

End for

For (j = 0 to P − 1) **do**

$PE_Max_VF(j) = \max_{1 \leq t \leq length(Jset(j))} Frqcy_matrix(Jset(t))$;

End

5.4 Necessary Optimal Voltage & Frequency Level Resource Collector (NOVFLRC)

This module is used to collect the information of PEs with optimal (minimum sufficient) voltage and frequency levels for the available PEs from the CRM. The module generates a new set of PEs (New_PE_Set) which are eligible enough to be replaced without affecting the performance but with reduced cost of the schedule. It begins by taking input PE_Max_VF (output produced by module MVFC); having record of maximum level of voltage and frequency of $Pseudo_Schedule$. Next, it sorts the $Avail_PE$ in ascending order of their maximum operating voltage and frequency levels in the variable $Sorted_Avail_PE$. It produces its output i.e., $New_{PE_{Set}}$ a set of PEs which are best fit (in terms of operating voltage and frequency levels) and highly eligible to generate a cost-effective schedule.

ACCEPTED MANUSCRIPT

Module 4: NOVFLRC

Input:

PE_{set}, F_{max}, PE_Max_VF;

Avail_PE; // *Record of available resources from Central Resource Manager*;

Output:

New_PE_Set; //*A set of PEs which are eligible for replacement*;

Begin

Sorted_Avail_PE = Sort(Avail_PE); //

// Sort the PEs in ascending order of their operating frequency & voltage;

For (j = 0 to P − 1) **do**

Flag = 0;

For (k = 0 to length(Avail_PE)) **do**

If ((PE_Max_VF(j) ≤ Sorted_Avail_PE(**k**))&& (Sorted_Avail_PE(**k**) < $fmax^j$)) **do**

// $fmax^j \in \mathbf{F_{max}}$;

New_PE_Set(j) = k;

Sorted_Avail_PE(**k**) = PE_Max_VF(j);

//Replace the kth PE of Avail_PE with j^{th} PE of Scheduled PE;

Flag = 1;

Break;

End if

If (flag == 0) **do**

New_PE_Set(k) = NULL; *// No suitable PE is found*;

End if

End for

End for

End

5.5 Price Calculator (PC)

This module fetches price list from pricing model (a component of the SP) which contains price rate ($ per time unit) for each PE in the cluster. Using this, it simply calculates the cost (*Pseudo_Total_Cost*) of the schedule generated by the PS. *Pseudo_Total_Cost* is also further required to calculate the gain for producing final cost-effective schedule.

ACCEPTED MANUSCRIPT

Module 5: PC

Input:

PE_{set}, makespan;

Rate_List;

//Collect from Pricing Model, contains price rate ($ per time unit) for each PE in the cluster;

Output:

Pseudo_Total_Cost; *//Total cost of the Pseudo Schedule*;

Begin

PE_Rate_List(j) = Fetch(Rate_List);

// Fetch the rate list for the used PEs which are in PE $(0 \leq j \leq P-1)$;

Calculate: Pseudo_Total_Cost $= \sum_{j=0}^{P-1}$ makespan × PE_Rate_List(j) ;

End

5.6 Resource Replacement Scheduler (RRS)

The RRS module is responsible to minimize the cost of the schedule as per the availability of the PEs.

Module 6: RRS

Input:

PE_{set}, Pseudo_Schedule, New_PE_Set ;

Output:

Final_Schedule: *//An energy efficient and cost effective final job schedule*;

Final_New_PE_Set; *//updated set of finally used PEs*;

Begin

For (j = 0 to P − 1) **do**

  If (New_PE_Set(j)! = NULL) **do**

    Replace(PE(j), New_PE_Set(j))

    // replace j^{th} PE with the PE contained by New_PE_Set(j) at same j^{th} index;

    Final_New_PE_Set(j) = PE_Set(j);

  Else

    Final_New_PE_Set(j) = PE(j);

  End if

End for

End

ACCEPTED MANUSCRIPT

This module begins by taking $Pseudo_Schedule$ and New_PE_Set as inputs. It replaces PEs of the $Pseudo_Schedule$ with the PEs contained in New_PE_Set. If New_PE_Set is NULL at the respective index it means, currently there is no PE available in CRM to replace the respective PE of the $Pseudo_Schedule$. Otherwise, replace the PEs of the $Pseudo_Schedule$ with the PEs available in New_PE_Set at the respective index. After completing this process, a final schedule is generated with the new set of PEs ($Final_New_PE_Set$).

5.7 Gain Calculator (GC)

The $makespan$, $Final_New_PE_Set$, $Pseudo_Total_Cost$ and $Rate_List$ are passed as inputs to the GC module. Where, $Rate_List$ has list of prices ($) for each PE offered by the pricing model. The $Pseudo_Total_Cost$ is the total cost ($) of the $Pseudo_Schedule$ which is calculated by PC. The objective of GC is to calculate % gain in the cost of the $Final_Schedule$ over $Pseudo_Schedule$ to measure the effectiveness of the model.

Module 7: GC

Input:

Makespan, Final_New_PE_Set, Pseudo_Total_Cost;

Rate_List; $//Collect\ from\ Pricing\ Model\ which\ contains\ price\ rate\ (\$\ per\ time\ unit)\ for\ each\ PE\ in\ the\ cluster$;

Output:

Gain; $// \%\ of\ obtained\ gain\ in\ comparison\ to\ Pseudo_Schedule$;

Begin

Final_Total_Cost = PC (Final_New_PE_Set, makespan, Rate_List);

Gain = (Pseudo_Total_Cost − Final_Total_Cost)/(Pseudo_Total_Cost) × 100;

End

These all input parameters are given to the PC (called as a function inside this module) and % of gain of the final schedule over $Pseudo_Schedule$ is calculated. Finally, EACEF offers a final green and cost effective schedule at the scheduler layer. Rest of the work is done by the SP by providing all necessary computing resources (on effective price) for the duration of the $makespan$ of the schedule. After successful execution, results are directly returned back to the client.

ACCEPTED MANUSCRIPT

6. Performance Evaluation by Simulation

The performance of the proposed EACEF is analyzed via simulation which verifies the effectiveness of EACEF in terms of % of saved energy within ***makespan*** extension limit as determined by green SLA. It is also to estimate the resource cost (pay per use) and % of the gain in the cost. In order to observe the behavior of EACEF in a truly heterogeneous environment, a mix heterogeneous workload generator [8, 37] is used to generate workload corresponding to eight different mathematical and computational problems (i.e., factorial, matrix multiplication, matrix chain multiplication, insertion sort, selection sort, bubble sort, matrix transposition and fibonacci series). The heterogeneous cluster is simulated with DVFS-enabled three real heterogeneous PEs [20] and their voltage and frequency pairs at respective levels, are shown in Table 2.

Table 2 Voltage and frequency pairs [20]

Level	PE^0		PE^1		PE^2	
	Voltage (V)	**Frequency(GHz)**	**Voltage (V)**	**Frequency(GHz)**	**Voltage (V)**	**Frequency(GHz)**
1	0.90	0.80	0.90	0.80	0.956	0.60
2	1.0	1.0	1.00	1.0	1.180	0.80
3	1.1	1.2	1.05	1.2	1.308	1.0
4	1.2	1.4	1.10	1.4	1.436	1.2
5	1.3	1.6	1.15	1.6	1.484	1.4
6	1.4	1.8	1.20	1.8	-	-
7	1.5	2.0	-	-	-	-

The PEs (PE^0, PE^1 and PE^2), as shown in Table 2, are AMD Athlon-64, AMD Turion MT-34 and Pentium M respectively. Further, the model is also tested using two well-known different schedulers i.e., HEFT [9] and ESHMP [10] which are designed to generate the schedule in the heterogeneous environment. The behavior of EACEF is observed on different performance parameters such as number of jobs, % of $makespan_extention$ (μ), Resource Requirement (RR) or the number of PEs and Resource Availability (RA). For all the experiments, the default testing parameters are shown in Table 3. The variation in these parameters, during experimentation, are discussed in the respective subsections. The % of saved energy, % of saved cost and the cost of the schedule is estimated using Eqs. 20, 21 and 22 respectively.

$$\%\ of\ saved\ Energy = \frac{EN\ by\ Base_Schedule - EN\ by\ Pseudo_Schedule}{EN\ by\ Schedule} \times 100 \quad (20)$$

$$\%\ of\ saved\ Cost = \frac{Cost\ of\ Pseudo_Schedule - Cost\ of\ Final_Schedule}{Cost\ of\ Pseudo_Schedule} \times 100 \quad (21)$$

ACCEPTED MANUSCRIPT

$$Cost\ of\ Schedule = \sum_{j=0}^{P-1} \frac{(makespan+\mu)\times Price_Rate^{j}}{10^{6}\times 60\times 60} \quad (22)$$

Table 3 Default testing parameters

Testing Parameters	Values
Number of jobs	1000
$makespan_extention\ (\mu)$	15%
Number of PEs (RR)	32
Resource Availability (RA)	25%

Where, "$EN\ by\ Base_Schedule$" denotes the energy consumed by the base schedule (HEFT/ESHMP). "$EN\ by\ Pseudo_Schedule$" is the energy consumption by the $Pseudo_Schedule$, an energy efficient schedule produced by the PS module of EACEF. Similarly, "$Cost\ of\ Pseudo_Schedule$" is the cost of $Pseudo_Schedule$ calculated using Eq. 22 and "$Cost\ of\ Final_Schedule$" is the cost of $Final_Schedule$ produced by RRS module of EACEF. To estimate the cost of the schedule, the rates of respective PEs are listed in Table 4. These are suggested by various cloud SPs for the number of similar resources [30, 32].

Table 4 Price rates of the PEs [20]

Processor IDs	Processor Names	Price_Rate ($/hour)
PE^{o}	AMD Athlon-64	0.095
PE^{1}	AMD Turion MT-34	0.068
PE^{2}	Intel Pentium M	0.046

Experiments are performed on randomly generated heterogeneous workloads as used in [8, 37]. Each experiment is executed 100 times and average for the final results are produced.

6.1 Varying Problem Size

The performance behavior of EACEF is analyzed on varying problem size (a number of jobs). The problem size is varied from 100 to 50000 jobs. In this set of experiments, except the problem size, other parameters (RR, μ and RA) are set to their default values (Table 3). The problem size in the respective experiment is shown on X-axis whereas Y-axis shows the respective performance metrics. Initially, two schedules (using HEFT and ESHMP schedulers) are generated on varied problem sizes. Further, EACEF is applied on the generated schedules and results are produced.

ACCEPTED MANUSCRIPT

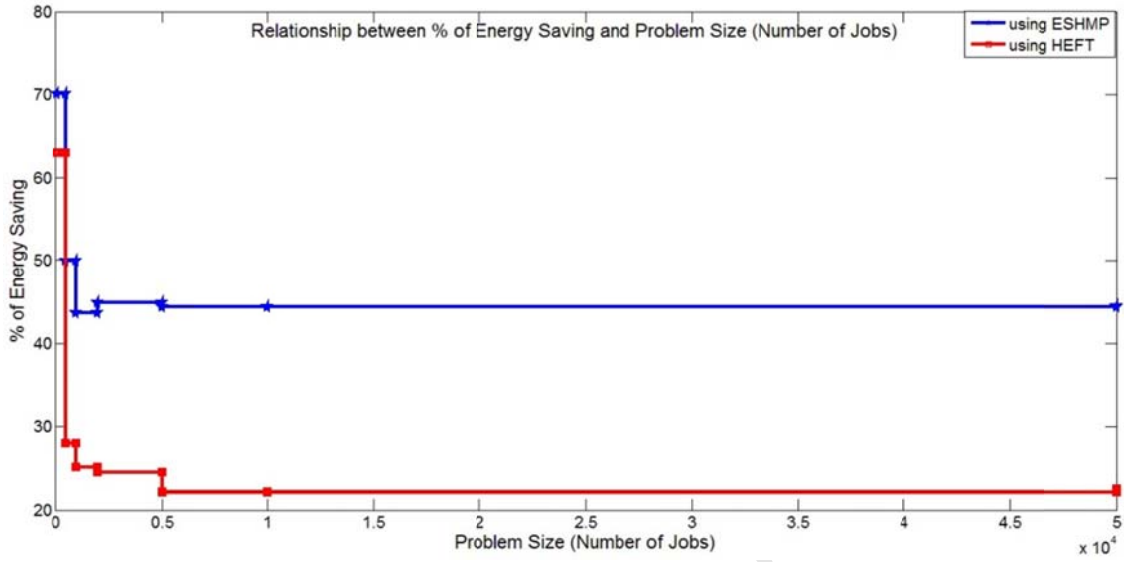


Fig. 2 Percentage of saved energy on varied problem size with HEFT and ESHMP schedules

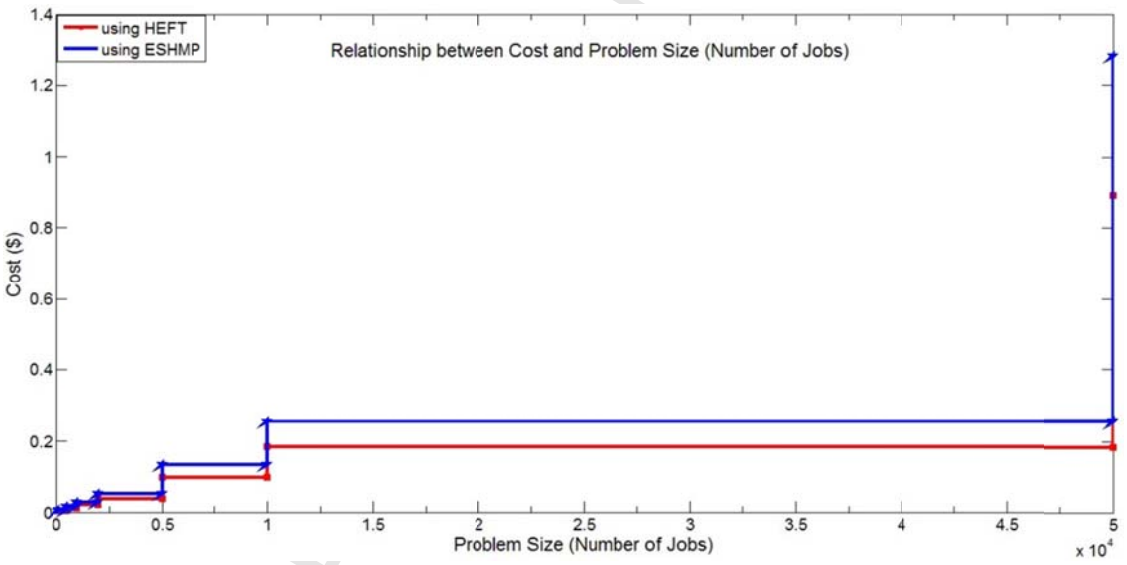


Fig. 3 Cost estimation on varied problem size with HEFT and ESHMP schedules

The results in Fig. 2, are used to observe the % of saved energy over two different heterogeneous schedulers i.e., HEFT and ESHMP. It is evident from the results that when the number of jobs are less, % of saved energy by both the schedulers is very high. On the other hand, when the number of jobs are increased, % of saved energy is low which becomes stable after a certain point. The reason being that initially the schedule has less number of jobs and large number of PEs of which all the PEs are not utilized properly and most of the PEs run on their threshold (minimum) voltage and frequency levels. It

results in low energy consumption. When number of jobs are increased to a limit at which all the PEs are utilized properly, then all the PEs run at the appropriate level of voltage and frequency and the computation is done within a bound of $makespan$. Hence, the energy consumption is little higher. Further, an increment in the number of jobs does not affect the computation and energy consumption ratio after settling on average load on each PE. The reason for this is μ being increased in the same ratio as per green SLA which provides the empty slacks to reduce the energy consumption in the same ratio. Thus, after a certain point at which sufficient amount of load is available to utilize all the PEs, the % of energy saving ratio is stable.

Fig. 3 shows the results corresponding to the cost of each schedule. The cost of the schedules generated by HEFT and ESHMP is estimated before the execution of RRS module of EACEF. From the results, it is observed that when the number of jobs are few, the cost of usage is lesser for both the schedules (generated by HEFT and ESHMP). Later, when the number of jobs are increased the cost of usage becomes higher. For the former case, the reason is that schedules had a few number of jobs and the PEs used for a small period. Since client plays a role only for the period ($makespan$) of usage of the resources, less number of jobs in the schedule with the same scheduler and resources (PEs) always have small $makespan$. As obvious, a small $makespan$ with fixed resources implies low cost ($). Similarly, when the number of jobs in the schedule are increased and the number of PEs are also fixed to the same default value (32), $makespan$ is obvious to increase. Hence the cost of resource usage by the schedules, also increases proportionally.

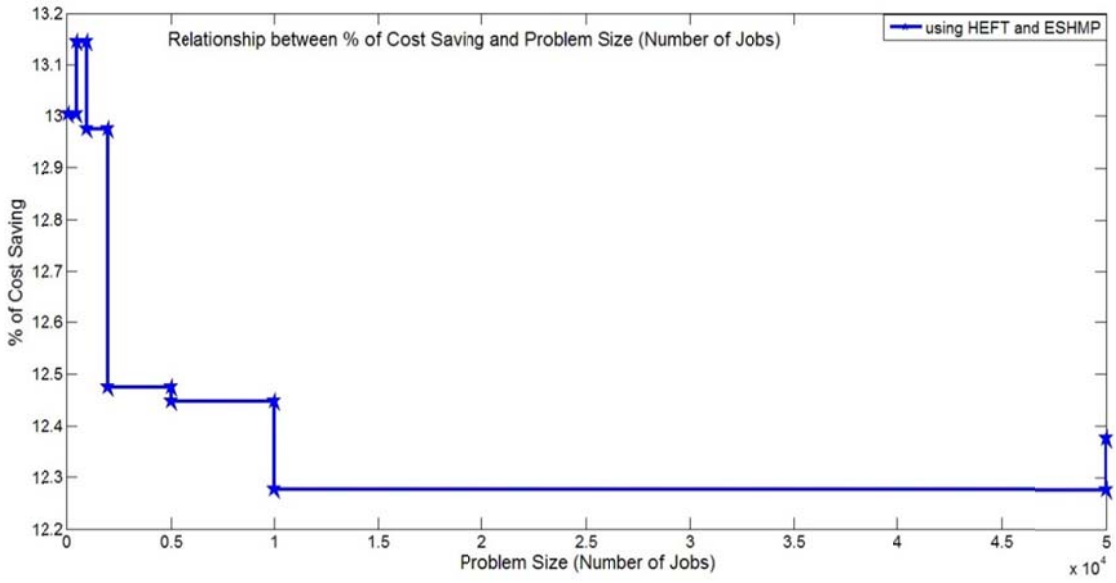


Fig. 4 A combined percentage of cost saving with varied problem size using HEFT and ESHMP schedulers

ACCEPTED MANUSCRIPT

Finally, Fig. 4 shows the estimated % of the cost saving for both the schedules. This is obtained after completing the execution of RRS module of EACEF. Note that this gain in the cost does not compromise the performance of the model. Moreover, the cost saving can also be enhanced if *RA* is higher. In this experiment, % of cost saving is estimated at *RA* of 25% fixed to its default value. From Fig. 4, it is observed that for the small problem size, % of cost saving is higher than the larger size of the problem. The reason is that cost is measured as a number of PEs slots used during the $makespan$. For small size of the problems, it is possible that some PEs are not used for the entire duration of the $makespan$ because the load is not enough to distribute all the PEs appropriately. So, there is little fluctuation in the % of cost saving. Later, it is also observed that for the large problem size, the % of cost saving becomes stable as enough jobs are available to distribute for all PEs appropriately.

With this set of experiments, it is observed that for the small size of the problem, energy and cost saving are higher because of less resource utilization. However, EACEF saved energy and cost for the large size of problems as well with a stable rate.

6.2 Varying Makespan Extension Rate

The performance behavior of EACEF is also analyzed by varying the % of $makespan$ extension (μ) as determined by the green SLA. The μ is varied between 5% to 30% and other parameters (problem size, *RR* and *RA*) are set to their default values. X-axis, of the respective Figs. 5, 6 and 7, displays the μ value and Y-axis shows the performance metric (% of saved energy/cost/ % of saved cost) for the respective experiments. Similar to section 6.1, model is verified on two schedules generated by HEFT and ESHMP schedulers and results are also collected in three parts.

Fig. 5 depicts the relationship between % of saved energy and μ. It is analyzed that when % of μ is increased, % of saved energy for the ESHMP schedule is slightly increased, whereas for HEFT schedule it slightly fluctuates. The reason is that the ESHMP schedule is not balanced enough as compared to HEFT schedule. So, when the value of μ is increased for the ESHMP schedule, it produces many empty slacks than HEFT schedule. More empty slacks indicate more energy saving for the schedule. It is also evident from the results that the HEFT schedule little fluctuates for different set of experiments. The reason is randomness in the heterogeneous workload.

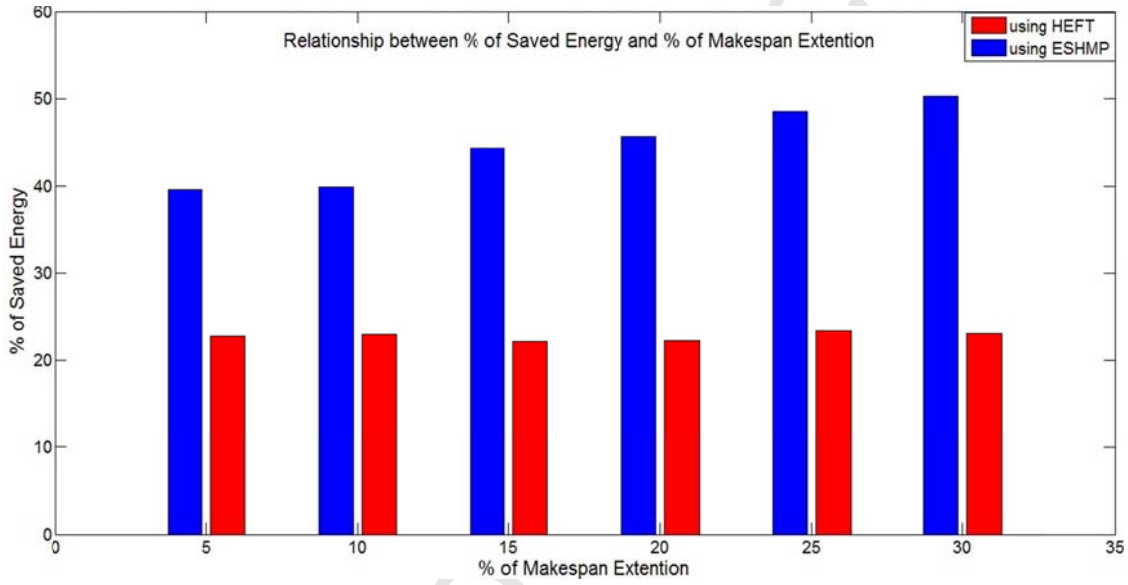


Fig. 5 Percentage of energy saving on varied makespan extension rate with HEFT and ESHMP schedules

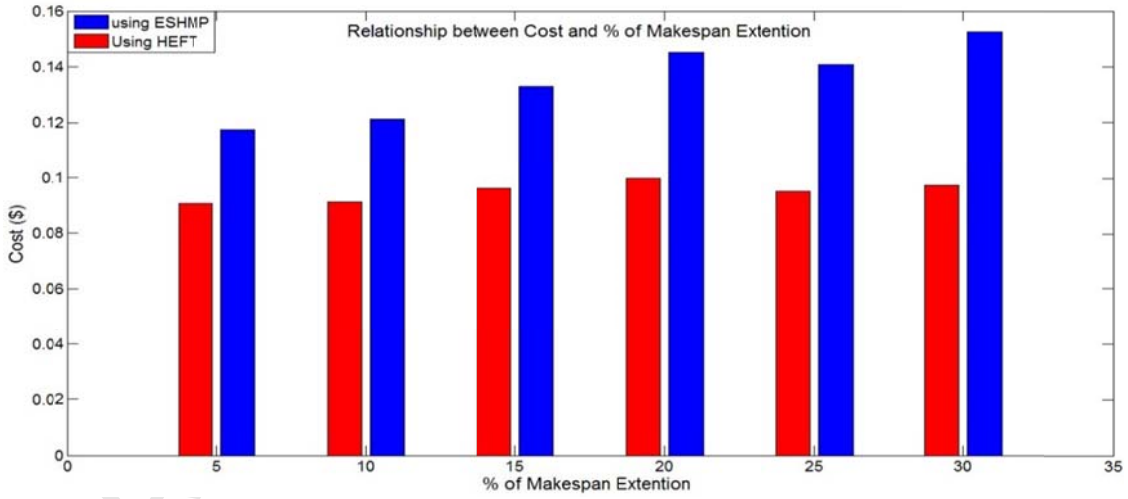


Fig. 6 Cost estimation on varied makespan extension rate for HEFT and ESHMP scheduler

ACCEPTED MANUSCRIPT

Fig. 6 shows the results for the second part of this experiment. It is to estimate the cost of the $Pseudo_Schedule$ generated by HEFT and ESHMP. It is analyzed that when the % of μ is increased, the cost ($) for ESHMP schedule is slightly increased and for HEFT schedule slightly swung. The reason is the same as discussed in the first part of this experiment, the schedule produced by the ESHMP scheduler is not balanced as compared to HEFT schedule. So, it is obvious that the $makespan$ of ESHMP is greater when the value of μ is increased. Moreover, $makespan$ due to empty slacks is also increased proportionally when % of μ is increased. Since PEs are used for the longer period of time for a greater $makespan$, this obviously results in higher cost. As mentioned, the HEFT schedule is more balanced. So, the increment in % of μ for HEFT schedule results in a proportional increment in the $makespan$ only, without extending extra empty slacks. As a result, there is no significant changes in the cost of HEFT schedule when % of μ is increased.

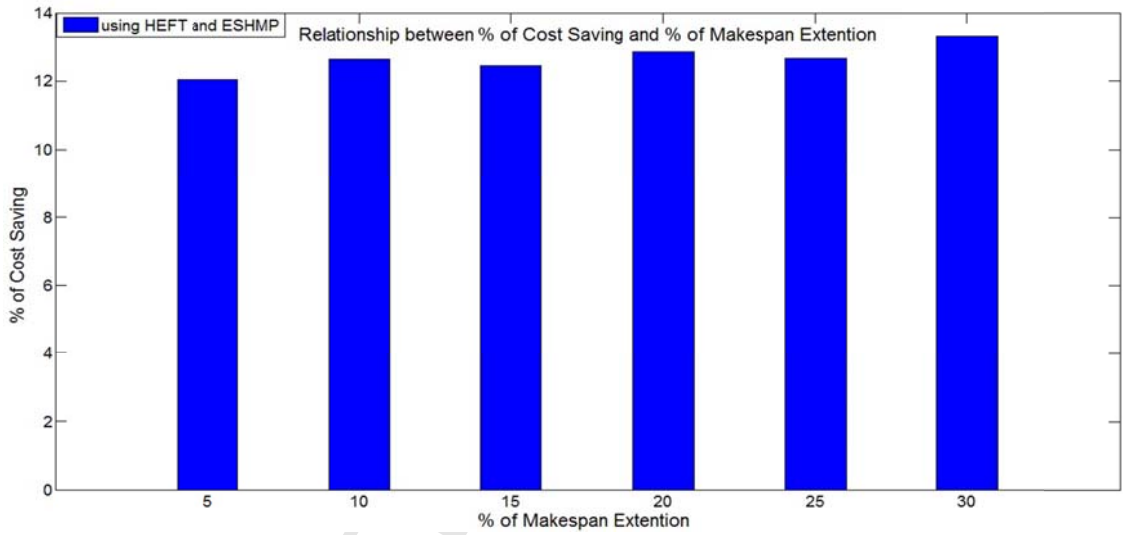


Fig. 7 A combined percentage of cost saving on varied makespan extension rate using HEFT and ESHMP schedules

Finally, % of gain in the cost of both the schedules is estimated in the third part of this experiment and results are shown in Fig.7. From the second part of this experiment, it is observed that the cost for both the schedules was different. But, in this part it is observed that the % of cost saving for both the schedules is same. It is because *RA* is same for both the schedules and the number of resources replaced for both the schedules using RRS module, are also same. Hence, despite of variation in the cost of the schedules, the % of cost saving with respect to their pseudo schedules is always same.

Thus, from this set of experiments, it is concluded that when μ is increased, the unbalanced schedule has the high rate of % of energy saving and higher cost for the $Pseudo_Schedule$ as compared to a balanced schedule. But due to same RA of both the schedules, the % of cost saving for the different schedules is unaffected. However, EACEF shows its suitability to achieve its goals while varying in μ.

6.3 Varying Resource Requirement

A set of experiments is also conducted to test the scalability of the model when RR (number of PEs) is varied from 16 to 512 while other parameters are set to their default values. X-axis of the respective Figs. 8, 9 and 10 depicts adjustment in the RR and Y-axis shows the performance metric (% of saved energy/ cost/ % of saved cost) for the respective experiments. Again, the model is tested for the two schedules generated using HEFT and ESHMP schedulers. Results are also collected in three parts as shown in Figs. 8, 9 and 10 respectively.

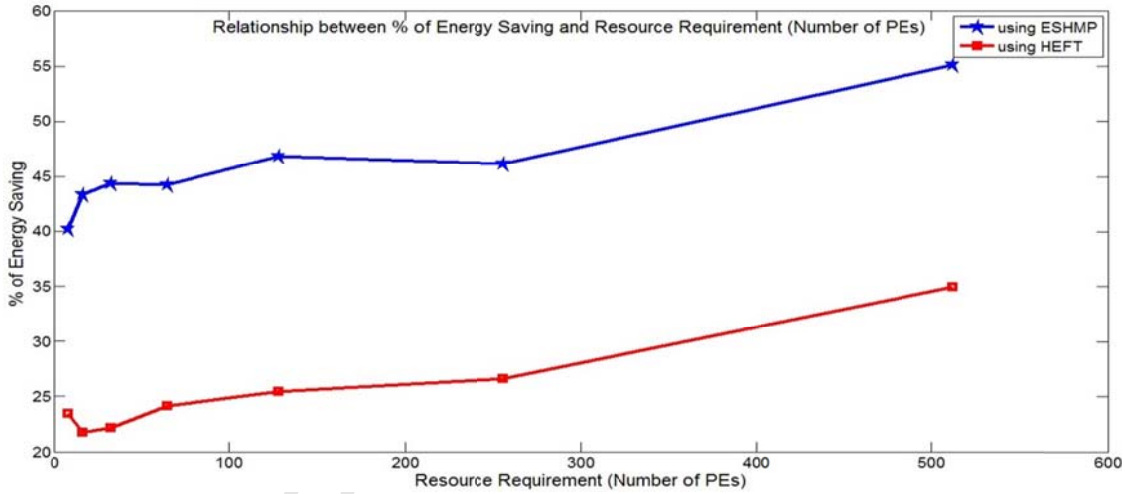


Fig. 8 Percentage of energy saving on varied RR with HEFT and ESHMP schedules

From Fig. 8, it is observed that the % of saved energy for both the schedules is slightly increased in a linear fashion (with a little fluctuation) when the number of PEs are increasing. It shows that the model preserves scalability characteristic. Also, when RR is very less and the number of jobs are relatively larger, the number of empty slacks, on the used PEs, are limited. Since % of energy saving is directly proportional to the empty slacks, the % of energy saving is less. On the other hand, when the number of PEs are increased then there is high probability of increment in the number of empty slacks which indicates more energy saving. Further, it is observed that % of energy saving for ESHMP is higher

as compared to HEFT. The reason is unbalanced nature of ESHMP schedule than HEFT schedule. A more balanced schedule has a low possibility of energy saving in comparison to the unbalanced schedule, (discussed in section 6.2).

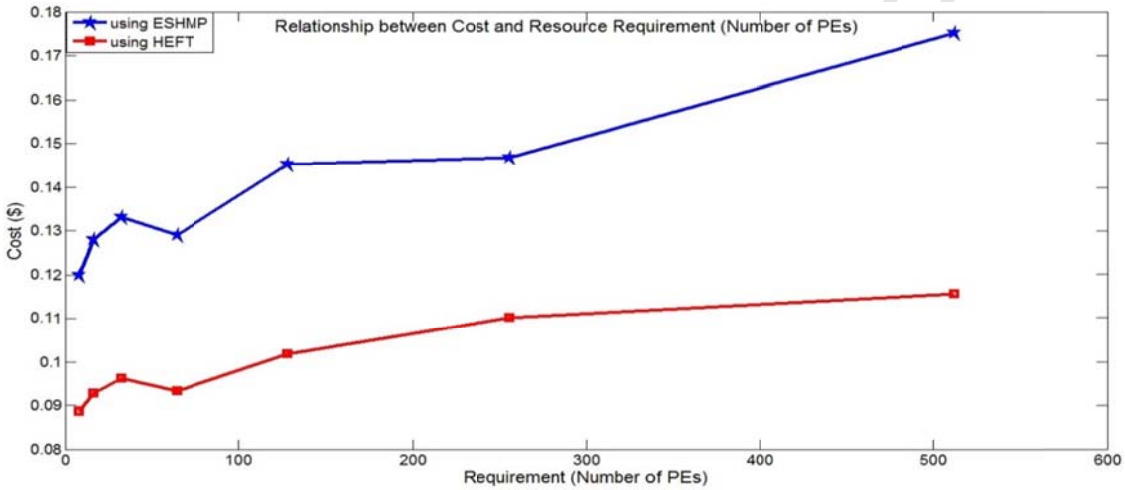


Fig. 9 Cost estimation on varied RR rate with HEFT and ESHMP schedules

Next, the cost for the pseudo schedules generated by HEFT and ESHMP is also estimated when *RR* is varied. From Fig. 9, it is clear that when the number of PEs are increasing the cost ($) for both the schedules is also slightly increased since the cost is directly proportional to *RR*.

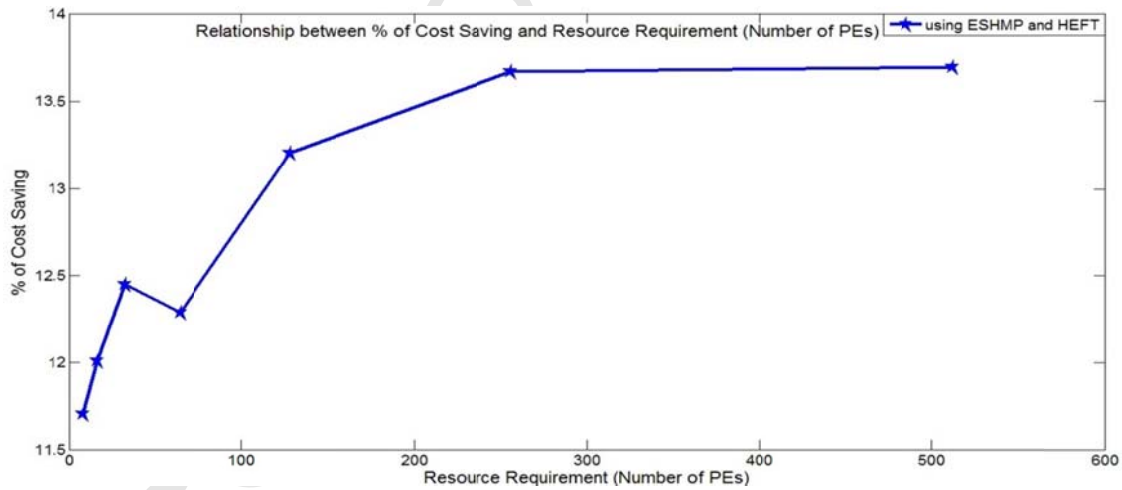


Fig. 10 A combined percentage of cost saving on varied RR using HEFT and ESHMP schedules

At the end of this experiment, the results (shown in Fig. 10) present an estimation in the % of cost saving for both the schedules. From Fig. 10, it is easy to verify that EACEF is gaining a noticeable

amount of % of cost saving (11.7% to 13.6%) despite of varied *RR* and limited *RA* (*RA*=25%). It is noted that increase in *RA* results in more gain in the % of cost saving. In brief, the scalability of EACEF is also verified in terms of both % of energy and cost saving, when *RR* is varied from 16 to 512 number of PEs.

6.4 Varying Resource Requirement and Resource Availability

In this subsection, the performance behavior of EACEF is analyzed by varying two different parameters i.e., *RR* and *RA* together. The relation between quantitative values of *RR* and *RA* is reserved for the respective experiments as shown in Table 5.

Table 5 Resource requirements and respective resource availabilities

RR(number of PEs)	16	32	64	128	256	512
RA(%)	15	20	25	30	40	50

The objective of conducting this set of experiments is to test the behavior of EACEF in terms of % of energy and cost saving when *RR,* and *RA* are varied together. To perform this set of experiments, number of jobs and μ are fixed to their default values.

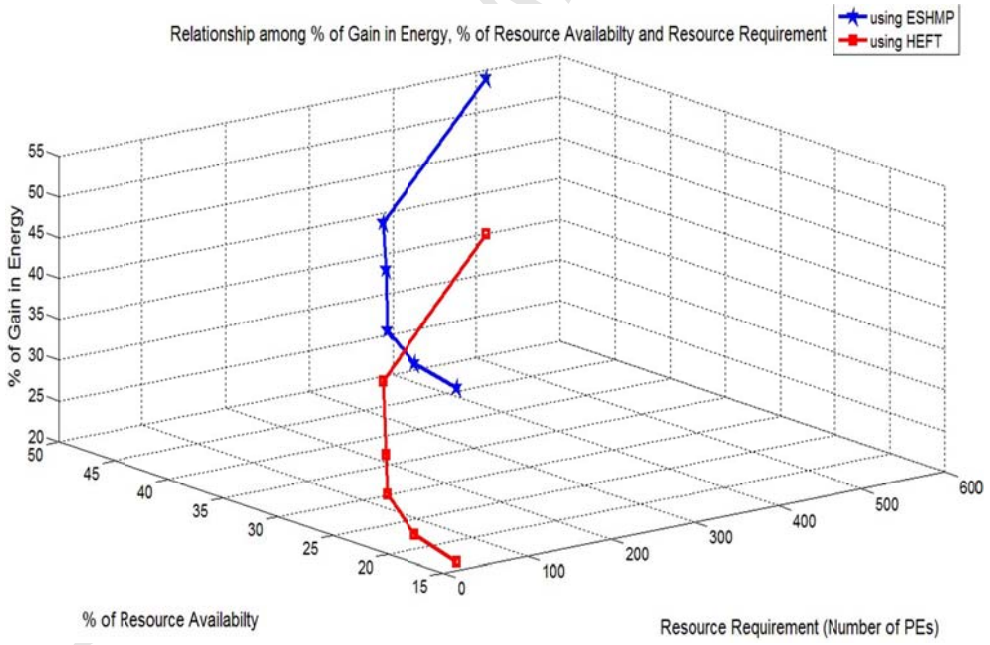


Fig. 11 Percentage of energy saving on varied RR and RA with HEFT and ESHMP schedules

Fig. 11 shows the results for the % of saved energy on two different schedules generated by HEFT and ESHMP. From the results, it is examined that EACEF is able to save upto 35% energy for

HEFT schedule and 55% energy for the ESHMP schedules with 50% available resources. Moreover, it is also observed that when *RR* (number of PEs) is high, model works more effectively. This proves that EACEF is highly scalable. It is also analyzed that there is a noticeable difference between % of saved energy between HEFT and ESHMP schedules. It is due to the different nature of the schedulers.

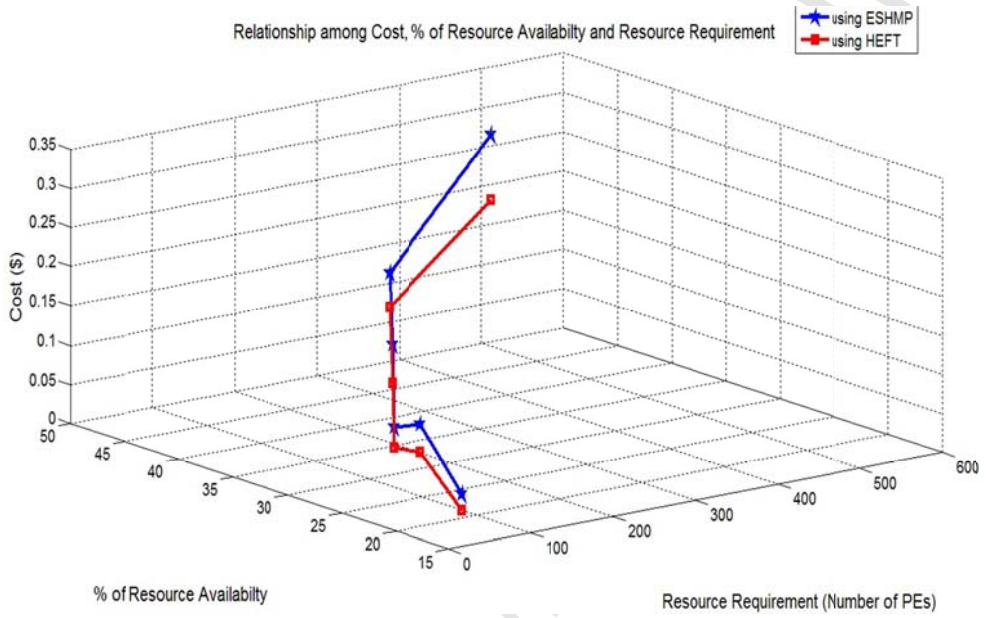


Fig. 12 Cost estimation on varied RR and RA rate with HEFT and ESHMP schedules

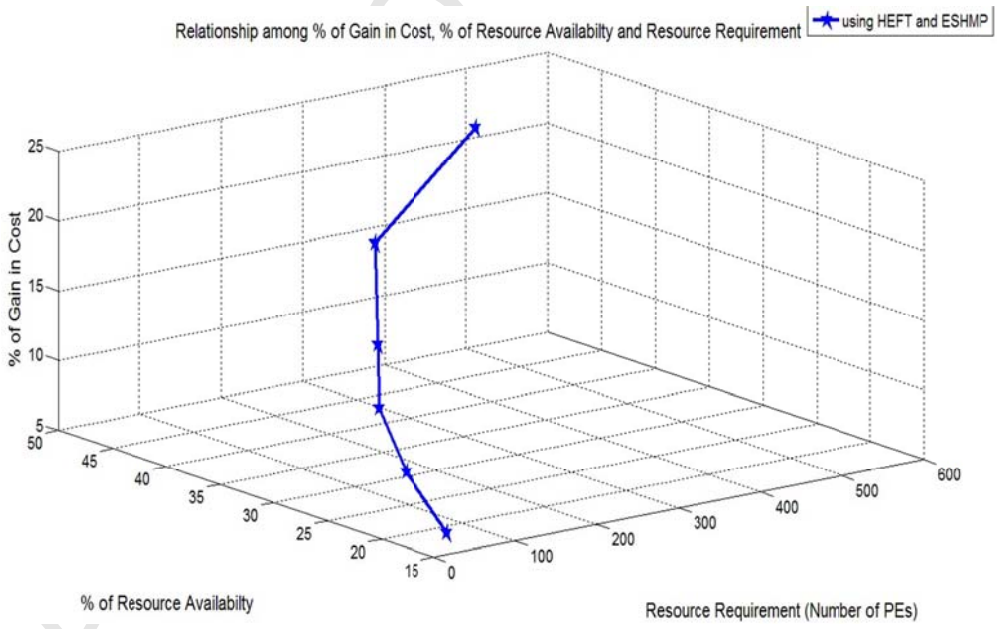


Fig. 13 A combined percentage of cost saving on varied RR and RA using HEFT and ESHMP schedules

ACCEPTED MANUSCRIPT

Further, the cost of the schedules generated by HEFT and ESHMP is also estimated and results are shown in Fig. 12. It is observed that the cost of both the schedules is same at some points and also differs at some other points on the plot. The reason is that some of the schedules generated by both the schedulers are relatively similar, and they occupy the resources in a similar fashion. So, the cost is nearly same at certain points on the plot. On the other hand, some of the schedules generated by HEFT and ESHMP are different in nature and they occupy the resources in a dissimilar fashion. So, the cost of both the schedules at some points differs.

Finally, it will be interesting to estimate % of the cost saving by EACEF when *RR* and *RA* are increased. From the results, shown in Fig. 13, it is observed that the % of saving in the cost increases significantly when both *RR* and *RA* are increased together, in comparison to only increment in *RR*. It is because if both (*RR* and *RA*) are increased, the number of required PEs and number of PEs are available in big quantity for replacement (as required by the RRS module). So, there are lot of choices to replace the PEs while executing RRS module. It results in more cost saving. With this experimental results, the scalability of EACEF is confirmed.

6.5 A Comparative Study

In this section, the performance of EACEF is compared with state of the art to study its effectiveness in terms of normalized energy consumption [33]. It is compared with two algorithms: Energy Aware Heterogeneous-Earliest-Finish-Time (EAHEFT) [34] and Enhanced Energy-efficient Scheduling (EES) [35]. Initially, HEFT has proven to perform well for task scheduling without considering energy [9, 36]. Later, energy aware HEFT is also introduced using DVFS technique in [35, 36]. EES is an enhanced energy-efficient scheduler based on DVFS technique. EES is used to decrease the frequencies of the jobs in a global manner and reassigns the tasks to appropriate time slots to reduce energy consumption.

ACCEPTED MANUSCRIPT

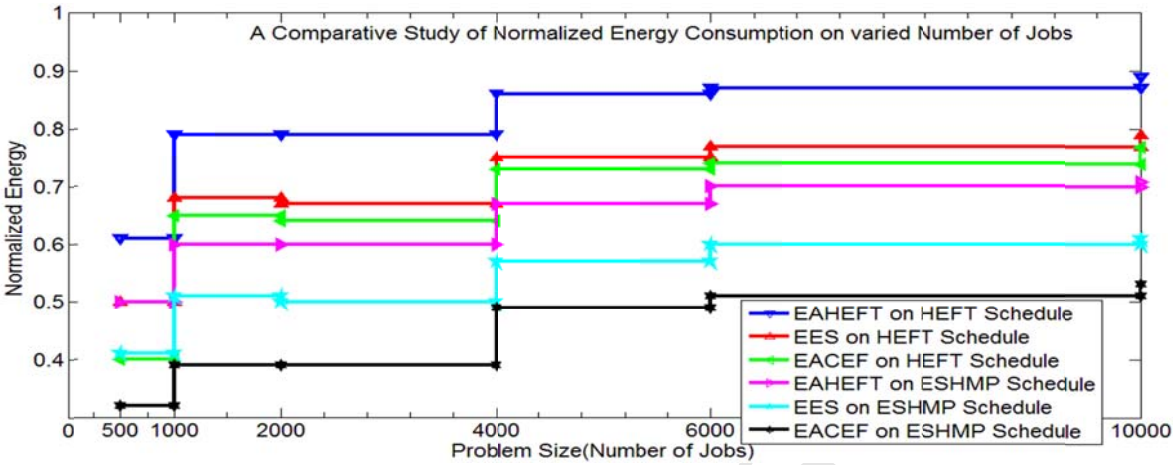


Fig. 14 A normalized energy consumption on varied number of jobs for HEFT and ESHMP schedules

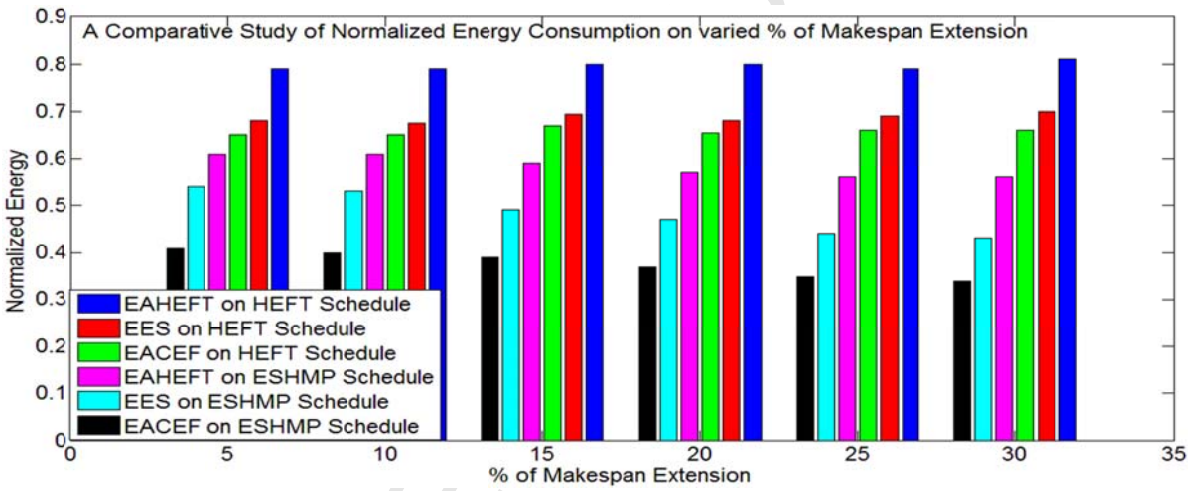


Fig. 15 Normalized energy consumption on varied % of μ for HEFT and ESHMP schedules

The comparative study of the algorithms is done for HEFT and ESHMP schedules on varied input parameters. The input parameters for the respective experiments are shown in respective Figs. while other parameters are default one. Normalized energy consumption is considered as a performance metric for this set of experiments [33]. A model which offers lower normalized energy consumption, is always preferred. Figs. 14, 15 and 16 plot the normalized energy consumption of three different models on varied number of jobs, % of *makespan* extension and number of PEs respectively. The models (EAHEFT, EES and EACEF) are also applied on both the schedules i.e., HEFT and ESHMP.

ACCEPTED MANUSCRIPT

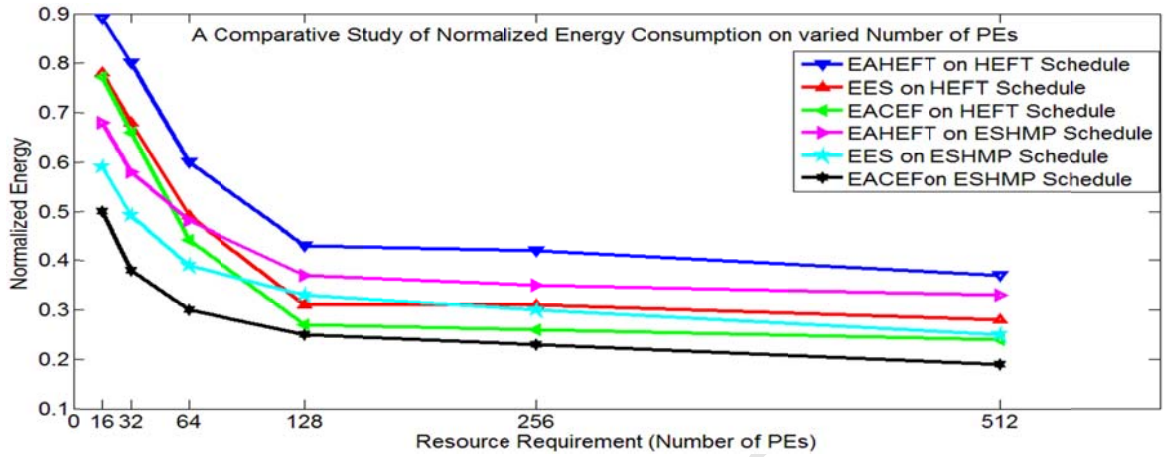


Fig. 16 Normalized energy consumption on varied number of PEs for HEFT and ESHMP schedules

From the Fig.14, it is observed that EACEF gains in its performance in comparison to EAHEFT and EES for both the schedules: HEFT and ESHMP, on varied number of jobs [500 to 10000]. EACEF consumes [0.42 to 0.77] and [0.32 to 0.53] normalized energy for the HEFT and ESHMP schedules respectively. Comparatively, EAHEFT and EES consume [0.61 to 0.89] and [0.5 to 0.79] for HEFT and [0.52 to 0.71] and [0.41 to 0.61] for ESHMP schedule respectively. Further, EACEF also minimizes its energy consumption in comparison to EAHEFT and EES, despite of variation in % of μ [5% to 30%] as shown in Fig. 15. As a summary, EACEF consumes [0.65 to 0.66] and [0.41 to 0.34] normalized energy for the HEFT and ESHMP schedules respectively when μ is varied. On the other hand, EAHEFT and EES consume [0.79 to 0.81] and [0.68 to 0.7] for HEFT and [0.61 to 0.56] and [0.54 to 0.43] for ESHMP schedule respectively. Furthermore, the scalability of the model is also sustained in more effective way as inferred from Fig. 16. EACEF consumes [0.77 to 0.24] and [0.5 to 0.19] normalized energy for the HEFT and ESHMP schedules respectively when number of PEs are varied [16 to 512]. EAHEFT and EES consume [0.89 to 0.37] and [0.78 to 0.38] for HEFT and [0.68 to 0.33] and [0.59 to 0.25] for ESHMP schedule respectively.

The reason for the significant improvement in the performance of EACEF is that it utilizes all possible idle slack rooms of the schedule and extended makespan (due to green SLA) to minimize energy consumption. To do this, DVFS technique is applied to scale down the voltage and frequency level of the PEs used in scheduling. Moreover, it also provides an opportunity to replace the high cost PEs with the low cost PEs, resulting in a novel, energy efficient and cost effective schedule.

7. Conclusion

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

This work presents a novel, automated, energy aware and cost effective resource assignment framework in heterogeneous cluster computing system. The objective is to assist the service providers to attract more clients by providing best services, despite of huge competition in the market. An advantage of this model is that it doesn't change the base schedule. It is capable of working on any existing base scheduler that is already being used for scheduling heterogeneous jobs and resources. For this reason, the model is verified on two different nature of heterogeneous base schedulers i.e., HEFT and ESHMP.

A model EACEF is proposed which works as a mediator between client and SP with the green SLA. Primarily, it operates on a base schedule (mappings between jobs and resources) generated by a base scheduler. An agreement is done for negotiated $makespan$ extension rate, energy and cost of the schedule and approved by the client during SLA. As mentioned, the model is dedicated to minimize energy consumption and cost ($) of the schedule. First, energy consumption is minimized using DVFS technique. Since DVFS technique has some energy and performance tradeoff to make it well-adjusted, empty slacks of the schedule on all the allocated PEs with extended $makespan$ are utilized to minimize energy consumption. Further, it is also observed that most of the fast PEs (with higher voltage and frequency levels) are not utilized properly. Using DVFS technique, these PEs run at their low voltage and frequency levels but as these PEs are very costly the RRS module of EACEF plays an important role to replace high-cost PEs with low-cost PEs (having sufficient voltage and frequency levels for the computation). This results in reducing the cost of the schedule without affecting the performance. Simulation study on EACEF is also done and results are compared with EAHEFT [34] and EES [35] state of the art. In summary, it is observed that EACEF comparatively saves a noticeable amount of energy and computation cost ($) despite of variations in different testing parameters. This model can be significantly useful in the competitive market to attract a number of clients to provide QoS on cheaper rates.

As the future work, it is proposed to extend the model in order to achieve the same objectives on dependency preserving parallel applications. In its current form, the model may not be efficient to apply on dependency preserving parallel applications.

References

[1] Maurizio Giacobbe, Antonio Celesti, Maria Fazio, Massimo Villari, Antonio Puliafito, Towards energy management in Cloud federation: A survey in the perspective of future sustainable and cost-saving strategies, Computer Networks, 91 (2015) 438–452.

[2] J. M. Kaplan, W. Forest, N. Kindler, Revolutionizing data center energy efficiency, technical report, Mckinsey and company, 2008,

ACCEPTED MANUSCRIPT

[3] L. Wang, S. U. Khan, D. Chen, J. Kołodziej, R. Ranjan, C.-z. Xu, A. Zomaya, Energy-aware parallel job model in a cluster, Future Generation Computer Systems, 29 (7) (2013) 1661–1670.

[4] D. Gmach, Y. Chen, A. Shah, J. Rolia, C. Bash, T. Christian, R. Sharma, Profiling sustainability of data centers, in: Sustainable Systems and Technology (ISSST), 2010 IEEE International Symposium on, IEEE, 2010, pp. 1–6.

[5] Y. C. Lee, A. Y. Zomaya, Energy conscious model for distributed computing systems under different operating conditions, Parallel and Distributed Systems, IEEE Transactions on 22 (8) (2011) 1374–1381.

[6] X. Zhu, C. He, K. Li, X. Qin, Adaptive energy-efficient model for real-time jobs on dvs-enabled heterogeneous clusters, Journal of parallel and distributed computing, 72 (6) (2012) 751–763.

[7] S. Baskiyar, R. Abdel-Kader, Energy-aware dag model on heterogeneous systems, Cluster Computing 13 (4) (2010) 373–383.

[8] Neetesh Kumar, D.P. Vidyarthi, "Improved Scheduler for Multi-Core Many–Core Systems," Journal of Computing. Springer, November 2014, Volume 96, Issue 11, 1087-1110.

[9] Topcuouglu, H., Hariri, S., Wu, M.Y.: Performance effective and low-complexity job model for heterogeneous computing. IEEE Trans. Parallel Distrib. System, 13(3), 2(2002)60–274

[10] Nie P, Duan Z Efficient and scalable scheduling for performance heterogeneous multicore systems. J. Parallel Distrib. Comput. (2012) 72:353–361.

[11] C.-M. Wu, R.-S. Chang, H.-Y. Chan, A green energy-efficient model using the dvfs technique for cloud datacenters, Future Generation Computer Systems, Vol. 37, July 2014, pp. 141–147.

[12] K. H. Kim, R. Buyya, J. Kim, Energy aware model of bag-of-jobs applications with deadline constraints on dvs-enabled clusters, in: CCGRID, Vol. 7, 2007, pp. 541–548.

[13] Y. Gao, H. Guan, Z. Qi, T. Song, F. Huan, L. Liu, Service level agreement based energy-efficient resource management in cloud data centers, Computers & Electrical Engineering, 40(5), July 2014, pp. 1621–1633

[14] M. Marzolla, R. Mirandola, Dynamic energy management for qos-aware applications, Sustainable Computing: Informatics and Systems 3 (4) (2013) 231–248.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

[15] M. E. Haque, K. Le, ´I. Goiri, R. Bianchini, T. D. Nguyen, Providing green slas in high-performance computing clouds, in: Green Computing Conference (IGCC), 2013 International, IEEE, 2013, pp. 1–11.

[16] A. Beloglazov, J. Abawajy, R. Buyya, Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing, Future Generation Computer Systems 28 (5) (2012) 755–768.

[17] Georges Da Costa, Marcos Dias de Assunção, Jean-Patrick Gelas, Yiannis Georgiou, Laurent Lefèvre, Anne-Cécile Orgerie, Jean-Marc Pierson, Olivier Richard, Amal Sayah, Multi-facet approach to reduce energy consumption in clouds and grids: the green-net framework, in: Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking, e- Energy'10, ACM, New York, NY, USA, 2010, pp. 95–104.

[18] Damien Borgetto, Michael Maurer, Georges Da-Costa, Jean-Marc Pierson, Ivona Brandic, Energy-efficient and SLA-aware management of IaaS clouds, in: Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet, e-Energy'12, ACM, New York, NY, USA, 2012, pp. 25:1–25:10.

[19] Jing Mei a, Kenli Li a, Jingtong Hub, Shu Yin a, Edwin H.-M. Shab,c, Energy-aware preemptive model for sporadic jobs on DVS platform, Microcores and Microsystems 37 (2013) 99–112.

[20] Huang, Q., Su, S., Li, J., Xu, P., Shuang, K., Huang, X. Enhanced energy-efficient model for parallel applications in cloud. In: 12th IEEE/ACM International Symposium Cluster, Cloud and Grid Computing (CCGrid), 2012, (2012) 781–786

[21] Fernando Koch, Marcos D. Assunção, Carlos Cardonha, Marco A.S. Netto, Optimising resource costs of cloud computing for education, Future Generation Computer Systems, 55 (2016) 473–479.

[22] Debdeep Paul, Wen-De Zhong, Sanjay K. Bose, Energy efficiency aware load distribution and electricity cost volatility control for cloud SPs, Journal of Network and Computer Applications 59 (2016) 185–197.

[23] Debdeep Paul, Wen-De Zhong, Sanjay K. Bose, Energy efficient cloud service pricing: A two-timescale optimization approach, Journal of Network and Computer Applications 64 (2016) 98–112.

[24] K. Le, R. Bianchini, J. Zhang,Y. Jaluria, J. Meng,T. D. Nguyen, Reducing electricity cost through virtual machine placement in high performance computing clouds, in: SC, 2011, p.22.

ACCEPTED MANUSCRIPT

[25] N. Bila, E. de Lara, M. Hiltunen, K. Joshi, H. A. Lagar Cavilla, M. Satyanarayanan,The case for energy oriented partial desktop migration, in: Proceedings of the 2Nd USENIX Conference on Hot Topicsin Cloud Computing, in:Hot Cloud'10, USENIX Association, Berkeley, CA, USA, 2010, p.3.

[26] I. n. Goiri, W. Katsak, K. Le, T. D. Nguyen, R. Bianchini, Parasol and greens witch: Managing data centers energyed by renewable energy, SIGPLAN Not. 48 (4) (2013) 51–64.

[27] H. Li, C. Wu, Z. Li, F. C. M. Lau, Profit maximizing virtual machine trading in a federation of selfish clouds, in: INFOCOM, 2013, pp.25–29.

[28] K. Le, R. Bianchini, J. Zhang, Y. Jaluria, J. Meng, T. D. Nguyen, Reducing electricity cost through virtual machine placement in high performance computing clouds, SC Conf. 0 (2011)1–12.

[29] R. Buyya, R. Ranjan, R. N. Calheiros, Inter cloud: utility oriented federation of cloud computing environments for scaling of application services, in: ICA3 PP (1), 2010, pp.13–31.

[30] Gaurav Baranwal, Deo Prakash Vidyarthi, A cloud service selection model using improved ranked voting method, Concurrency Computat.: Pract. Exper. Wiley Online Library (wileyonlinelibrary.com), (2016) DOI: 10.1002/cpe.3740.

[31] Kyong Hoon Kim, Rajkumar Buyya, Jong Kim, Energy aware Algorithm of bag of- tasks applications with deadline constraints on DVS-enabled clusters, in: CCGRID, 2007, pp. 541–548.

[32] http://blog.cloudharmony.com/2010/05/what-is-ecu-cpu-benchmarking-in-cloud.html

[33] Jing Mei a, Kenli Li a, Jingtong Hub, Shu Yin a, Edwin H.-M. Shab,c, Energy-aware preemptive scheduling model for sporadic tasks on DVS platform, Microprocessor and Microsystems 37 (2013) 99–112.

[34] Lee, Y.C., Zomaya, A.Y.: Energy conscious scheduling for distributed computing systems under different operating conditions. IEEE Trans. Parallel Distrib. Syst. 22, 1374– 1381 (2011).

[35] Huang, Q., Su, S., Li, J., Xu, P., Shuang, K., Huang, X.: Enhanced energy-efficient scheduling for parallel applications in cloud. In: 12th IEEE/ACM International Symposium Cluster, Cloud and Grid Computing (CCGrid), 2012, pp. 781–786 (2012)

ACCEPTED MANUSCRIPT

[36] Zhuo Tang, Ling Qi , Zhenzhen Cheng, Kenli Li , Samee U. Khan, Keqin Li An Energy-Efficient Task Scheduling Algorithm in DVFS-enabled Cloud Environment, J Grid Computing. 2015, DOI: 10.1007/s10723-015-9334-y.

[37] https://drive.google.com/open?id=0B7P0JfSxxK-wN1ozUWtFelVaeGs.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biographies (Text)**

Neetesh Kumar received his M.tech in computer science and technology from school of Computer and system Sciences of Jawaharlal Nehru University India. He has also received his PhD. from Jawaharlal Nehru University. Currently, he is working as an assistant professor at Shri Mata Vaishno Devi University, Jammu & Kashmir India. His research interest includes parallel and distributed computing, high performance computing.

Deo Prakash Vidyarthi, is Ph.D. in Computer Science. After serving Banaras Hindu University, Varanasi for more than 12 years, he joined JNU as Associate Professor in the year 2004 and currently working as Professor in the School of Computer & System Sciences, Jawaharlal Nehru University, New Delhi. Dr. Vidyarthi has published around 60 research papers in various International Journals and Transactions (including IEEE, Elsevier, Springer, Wiley, World Scientific etc.) and around 35 research papers in proceedings of various peer-reviewed conferences in India and abroad. He has contributed chapters in many edited books. He is in the editorial board of many International Journals and also in the reviewer's panel of many International Journals. Dr. Vidyarthi has co-authored a book (research monograph) entitled "Scheduling in Distributed Computing Systems: Design, Analysis and Models" published by Springer, USA released in 2009. Another book (edited) by Dr. Vidyarthi is "Technologies and Protocols for the Future Internet Design: Reinventing the Web", by IGI-Global (USA) released in the year 2012. Dr. Vidyarthi is the member of the IEEE, International Society of Research in Science and Technology (ISRST), USA, senior member of the International Association of Computer Science and Information Technology (IACSIT), Singapore and International Association of Engineers (IAENG). His research interest includes Parallel and Distributed System, Grid and Cloud Computing, Mobile Computing.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

***Biographies (Photograph)**





ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Highlights

1. An Energy Aware Cost Effective Framework (EACEF) is proposed for scheduling in heterogeneous cluster environment.
2. The system model uses DVFS technique to minimize energy consumption with green SLA.
3. Resource replacement module is introduced to reduce the computation cost ($).
4. Model utilizes the heterogeneous characteristics of jobs and computing resources to minimize energy and cost.
5. Effectiveness of EACEF is verified through simulation using real heterogeneous PEs on two schedulers (HEFT and ESHMP).

ACCEPTED MANUSCRIPT

An Object-Oriented MDP Representation for Robotic planning

Aasheesh Singh, Department of Electronics & Communication, Delhi Technological University
(asjsingh97@gmail.com)

***Abstract*—This paper aims to implement Object-Oriented Markov Decision Process (OO-MDPs) for goal planning and navigation of robot in an indoor environment. We use the OO-MDP representation of the environment which is a natural way of modeling the environment based on objects and their interactions. The paper aims to extend the well known Taxi domain example which has been tested on grid world environment to robotics domain with larger state-spaces. For the purpose of this project we have created simulation of the environment and robot in ROS (Robot operating system) with Gazebo and Rviz as visualization tools. The mobile robot uses a 2D LIDAR module to perform SLAM in the unknown environment. The goal of this project is to be able to make an autonomous agent capable of performing planning and navigation in an indoor environment to deliver boxes (*passengers* in Taxi domain) placed at random locations to a particular location *(warehouse)*. The approach can be extended to a wide variety of mobile and manipulative robots**

***Index Terms*—Markov decision process, SLAM, Robot operating system, Reinforcement learning, Adaptive Monte-Carlo Localization.**

1. INTRODUCTION

Markov Decision process is a discrete time stochastic control process. At each time step, the agent is in some state **s,** and takes an action **a** which takes it to its next state **s'** giving the agent a reward $\mathbf{R_a(s,s')}$**.** The core problem of MDPs is to find a "policy" for the decision maker: a function $\boldsymbol{\pi}$ that the decision maker will choose when in state **s.** Algorithms for Reinforcement learning in MDP environments suffer from what is known as the curse of dimensionality: an exponential explosion in the total number of states as a function of the number of state variables. Learning in environments with extremely large state spaces is challenging if not infeasible without some form of generalization.

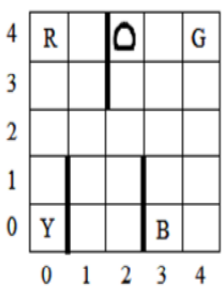
The OO-MDP representation starts from attributes that can be directly perceived by the agent, rather than predicates introduced by the designer, this similar kind of formalism has been used in Relational MDPs (RMDPs), introduced by Guestrin et al. (2003) in the context of planning. The *Taxi domain* example has been successfully tested with a videogame "*Pitfall"* in the paper "*An Object-Oriented Representation for Efficient Reinforcement Learning"* by Deuk, Littman & Cohen. We aim to extend this approach to mobile robots in an indoor grid like environment.

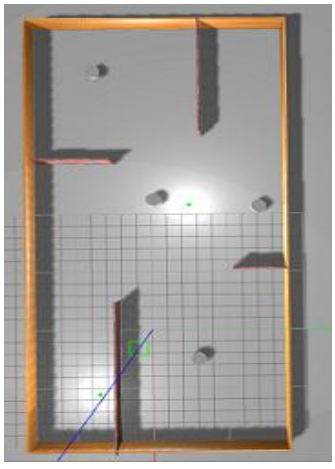
2. Notation

We use a standard Markov Decision Process (MDP) notation throughout this paper (Puterman, 1994). A finite MDP M is a five tuple <S,A,T,R,γ>. We use T(s'|s,a) to denote the transition probability of state s'given state– action pair(s,a)and R(s,a)to denote the expected reward value. A deterministic MDP is one in which there is a single next state s' for every given state s and action a; that is, ∀s ∈S, a∈ A, ∃s' ∈S: T(s'|s,a)=1.

3.Environment Representation

We will use the Taxi domain, defined by Dietterich (2000), as an example to introduce our formalism. Taxi is a grid world domain (see Figure 1.a), where a taxi has the task of picking up a passenger in one of a pre-designated set of locations (identified in the figure by the letters Y, G, R, B) and dropping it off at a goal destination, also one of the pre-designed locations. The set of actions for the taxi are North, South, East, West, PICKUP and DROPOFF. Walls in the grid limit the taxi's movements. Fig. 1.b shows the simulated indoor environment in Gazebo consisting of walls, obstacles and cylindrical boxes to be picked (*passengers*).





S.no	Class name	Attributes of class	No. of objects
1.	Agent	Odometry data	1
2.	Cylindrical Boxes(passenger)	Odometry, Boolean variable in_bot	Variable
3.	Walls/Obstacles	Odometry coordinates	Variable
4.	Destination	Odometry coordinates	1

The OO-MDP representation of the environment consisting of its objects is shown in the tabular form. The class Passenger has an additional attribute in.bot which returns true, when the object has been picked.

When two objects interact in some way, they define a **relation** between them. A combination of the relation established, plus the internal states of the two objects, determines an effect—a change in value of one or multiple attributes in either or both interacting objects. This is a *significant change in property* from classical MDPs which consider walls as a property of the specific location of the grid whereas in OO-MDP , the interaction(hence the relation) is same irrespective of the location. For our bot representation, we will define 5 relations: $touch_N(o1,o2)$, $touch_S(o1,o2)$, $touch_E(o1,o2)$, $touch_W(o1,o2)$ and on(o1,o2), which define whether an object o2 ∈ Cj is exactly one cell North, South, East or West of an object o1 ∈ Ci, or if both objects are overlapping (same x, y coordinates).

4. Data structure and definitions:

- **T** is the union of all terms t that will be involved in the conditions that determine the transition dynamics of the environment described by the OO-MDP, plus their negations **¬t**, with |T|=2n.
- **Cond(s)** function returns all the terms that are true in state s.
- An **effect E** is change of attributes required to cause a transition from state s to s' with action a.
- For comparing two conditions c1 & c2 we have the commutative operator defined as :

c_1	c_2	$c_1 \oplus c_2$
0	0	0
1	1	1
0	1	*
0|1	*	*

- For any states **s** and **s'** and attribute **att**, the function $\mathbf{eff_{att}(s,s')}$ returns one effect of each type that would transform attribute att in s into its value in s'.
- A prediction p is a pair **(p.model, p.effect)**, where p.model is a condition that represents the set of terms that need to be true for p.effect to occur.
- If for an action a, s=s' then that condition is called *failure condition*, denoted by $\mathbf{F_A}$.

- The relations together with the attributes define the state of the system, here we will have a 7 bit array describing the state of the system as: Suppose the bot is in the position (2,4) of the grid world defined in fig. 1then ***cond(s)*** returns:

  {$touch_N$(bot,wall), ¬ $touch_S$(bot,wall),
  ¬ $touch_E$(bot,wall), $touch_W$(bot,wall),
  ¬on(bot,passenger),¬on(bot,destination),
  passenger.in_bot = T }
  forming a **7 bit array= 1001001**.

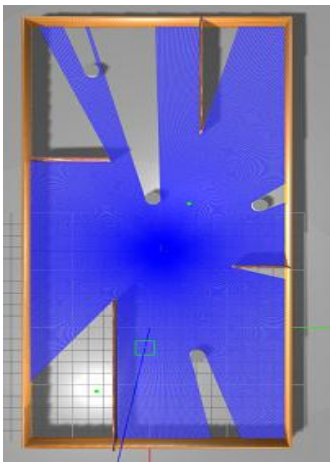


Fig. 3: Overhead view of the Laser graph when the bot is placed in the middle of the room. Passenger boxes & obstacle walls deflect the laser beam while it passes through the rest of the area(shown in blue). From here, we can form the touch(o1,02) relations between objects.

5. Mapping and localization

The robot uses **Adaptive Monte-Carlo localization** to localize itself in the environments given the map of the environment. In ROS the "gmapping" package generates the map of the environment given the laser scan data, the transform frame and odometry topics. The approach derives from earlier work on Markov localization which represents the robot's belief by a probability distribution over possible positions, and use Bayes rule and convolution to update the belief whenever the robot senses or moves. The idea of probabilistic state estimation goes back to Kalman filters which use multivariate Gaussians to represent the robot's belief. MCL uses fast sampling techniques to represent the robot's belief. When the robot moves or senses, importance re-sampling is applied to estimate the posterior distribution. An adaptive sampling scheme which determines the number of samples on-the-fly, is employed to trade-off computation and accuracy.

A. Advantages of using Monte carlo localization over Kalman filter and Markov localization:

- In contrast to existing Kalman filtering based techniques, it is able to represent multi-modal distributions and thus can globally localize a robot.
- It drastically reduces the amount of memory required compared to grid-based Markov localization and can integrate measurements at a considerably higher frequency.
- It is more accurate than Markov localization with a fixed cell size, as the state represented in the samples is not discretized.

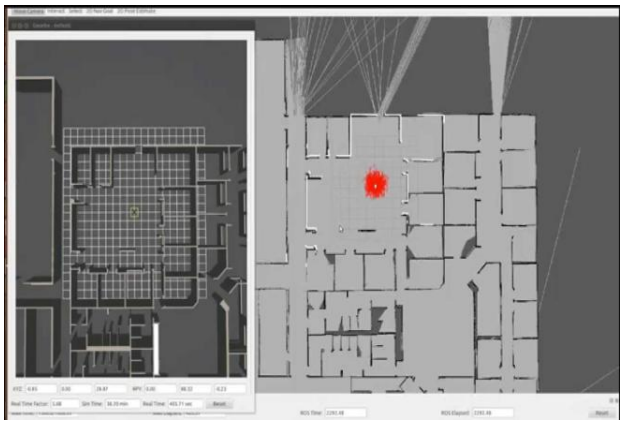


Fig. 3 AMCL simulation on a bot in a maze like environment. The red cloud around the bot shows the robot's belief/uncertainty in its pose while localization.

Amcl takes in a laser-based map, laser scans,and transform messages, and outputs pose estimates. On startup, amcl initializes its particle filter according to the parameters provided.

6. Discovering the optimal policy/Optimisation:

A. General transition dynamics:

For each state-action pair function **predictTransition(s,a)** returns a predicted next state s' according to *p.model*. The agent tries to discover the optimal policy $\boldsymbol{\pi}$ by running repetitive episodes of the same task to convergence.
The two main routines of the algorithm are predictTransition (Algorithm 1), which pre-dicts the next state given a current state and action based on the current model, and **addExperience** (Algorithm 2), which learns a model of the OO-MDP. If predictTransition is not able to predict a next state with accuracy, it returns $\boldsymbol{s_{max}}$.

We state the algorithm for the **predictTransition(s,a)** as follows:

Input: (s,a) **Output:** predicted state s'

1: if $\exists c \in F_a$ s.t. $c \models cond(s)$ **then**
// C is a known failure condition in the environment(**eg. Collision with walls and obstacles**)
Return s
2: **else**
3: **for all** attributes att ∈ $U_{c \in C}$ Att(c) **do**
4: E ←∅
5: **if** $\exists p \in pred(a, att, type)$ s.t. $p.model \models cond(s)_S$ **then**
6: Add *p*.effect to *E*
7: **end if**
8. **end for**
9: **if** $E = \varnothing \vee \exists e_i, e_j \in E$ s.t. e_i and e_j are incompatible **then**
10: Return s_{max}
11: **else**
12: // Set *E* contains all the individual operations that need to be applied to attributes in *s* in order to convert it to s'.
13: *s* ← apply *E* to *s*
14: Return s
15: **end if**
16: end if

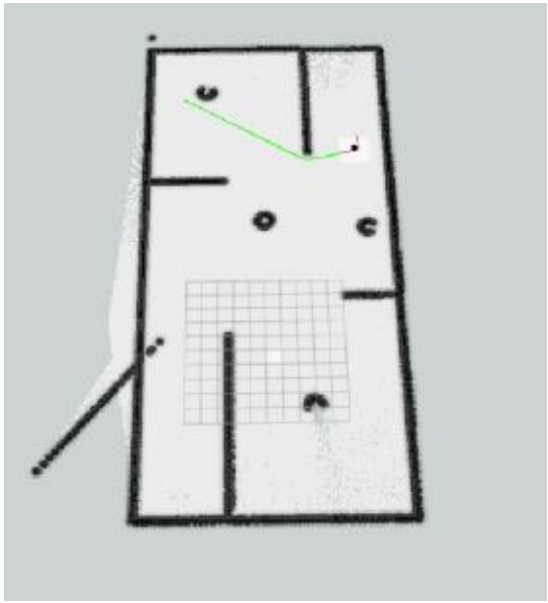


Fig. 4: Rviz simulated map of the grid like indoor environment. Gmapping node returns the map of the environment using the data published on the /sensor_mgs/LaserScan topic by the 2D LIDAR module. The green line indicates the predicted path(next state) of the bot as it advances.

We will now state a theorem to state the worst case bound i.e. the maximum no. of steps required to attain optimal policy so that we can state our next algorithm on **addExperience(`s,a,s',k`)**

Theorem: *The transition model for a given action a, at-tribute att and effect type type in a deterministic OO-MDP is KWIK-learnable with a upper bound of* $O(nk + k + 1)$, *where* n *is the*

number of terms in a condition and k is the maximum number of effects per action–attribute.

Algorithm 2 `addExperience(s,a,s',k)` method

Inputs: an observation $< s, a, s' >$,*k*;, the maximum number of different effects possible for any action, attribute and effect type.
if $s = s'$ **then**
  // Found a failure condition for action *a*, update F_a
  Remove all $c \in F_a$ s.t. $cond(s) \models c$.
  $F_a \leftarrow F_a \cup \{cond(s)\}$
else
  for all attributes $att \in \bigcup_{c \in C} Att(c)$ **do**
    for all $e \in eff_{att}(s, s')$ **do**
      Find a prediction $p \in pred(a, att, e.type)$ such that p.effect = e
      if $\exists p$ **then**
        // We already have a (condition, effect) prediction for current *a*, *att*, and *type*. Update condition and verify that there are no over-laps.
        $p.model \leftarrow p.model \oplus cond(s)_S$.
        if $\exists c \in (pred(a, att, e.type) \setminus p).models$ s.t. $p.model \models c$ **then**
          // Conditions overlap, violating an assumption, meaning it is not the right *type* of effect for this action and attribute.
          Remove $pred(a, att, e.type)$ from P
        end if
      else
        // We observed an effect for which we had no prediction. If its condition does not over-lap an existing condition, then add this new prediction.
        if $\exists c \in pred(a, att, e.type).models$ s.t. $cond(s) \models c \vee c \models cond(s)$ **then**
          Remove $pred(a, att, e.type)$ from P
        else
          Add $(cond(s), e)$ to $pred(a, att, e.type)$.
          // Verify that there aren't more than *k* predictions for this action, attribute and type.
          if $|pred(a, att, e.type)| > k$ **then**
            Remove $pred(a, att, e.type)$ from P
          end if
        end if
      end if
    end for
  end for
end if

7. Realization in Robot operating system(ROS)

We form a grid world like structure mirroring the Taxi-domain example by dividing our environment with grid size= footprint of the robot. The physical meaning of the various parameters as described in the taxi domain can be represented as:

- S= current x,y coordinate frames returned by odometry data and determining the pose using AMCL algorithm upto certain belief.
- S'= next estimated state for goal-planning

- The laser scan data returns the distance from a obstacle nearby which is divided into grid units and the therefore the relation touch(agent,wall) is formed which tells the bot of next possible actions. For eg: If $touch_N$(agent,wall) is true the bot cannot give velocity commands in the north/forward direction as it would lead to a failure state or collision.

- The effect type E gives all the possible change in attributes(x,y coordinates) that has to take place according to the policy.

- After every episode/simulation the robot returns data collected in the function pred(a,att,e.type) which is updated over the next episode and hence optimal policy to discover the shortest possible path to achieve the goal is discovered over various iterations. (Algorithm 2)

8. Conclusion and Future work:

MDP like environment suffer from what is known as the curse of dimensionality due to the exponential number of state variables as state-spaces increase. OO-MDP greatly simplifies this problem by reducing state-spaces by making MDP properties(location of walls, obstacles) as objects. But still the approach is currently valid for static transition dynamics as for stochastic environments a more complex learning algorithm would be needed to learn transitions effectively in the face of noise. We aim for extending this approach to much dynamic environments and hence a rather more effective learning algorithm for it.

9. Acknowledgment

I would like to thank Dr. Balaraman Ravindran, Dept. of Computer Science, IIT Madras under whose guidance i got the opportunity to do my first research project on Reinforcement learning. Also i would like to thank my mentors Priyatosh Mishra (Phd.) and Mohammed Hisham at RISE Lab, IIT Madras for guiding me at every step of the project and learning experience.

10. References

Diuk,Cohen,L.Littman "An object oriented representation for efficient reinforcement learning

Dietterich, T. G. (2000). Hierarchical reinforcement learning with the MAXQ value function decomposition. *Journal of Artificial Intelligence Research*, *13*, 227–303.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. The MIT Press.

Commuters' Exposure to $PM_{2.5}$: A Case Study of Delhi

Rajeev Kumar Mishra[1*], Amrit Kumar[2], Ankita Shukla[3]

Abstract Number: 16

[1*] Assistant Professor (Corresponding Author)
Department of Environmental Engineering,
Delhi Technological University,
Delhi -110 042, INDIA
Email: rajeevkumarmishra@dce.edu

[2]Research Scholar
Department of Environmental Engineering,
Delhi Technological University,
Delhi -110 042, INDIA
Email: amrit.kumar@dtu.co.in

[3]Environment Manager,
JK Infraprojects Limited,
CC-24, Delhi
Email: enviro1657@gmail.com

Abstract

This study is an effort to assess the exposure of particulate matter ($PM_{2.5}$) to commuters, travelling in different modes of transport (AC Bus, Non-AC Bus, AC Car, Non-AC Car and Auto-rickshaw) along major transport corridor in Delhi. The study was conducted during morning peak hours, evening peak hours and off-peak hours of the day. Among all considered modes of transport, the maximum concentration was found in Auto-rickshaw (102 µg/m^3), while the minimum was observed in AC car (43 µg/m^3) and AC bus (50 µg/m^3), which indicates that the public transport may be used by commuters to reduce their personal exposure to particulate matter.

Keywords: *Commuter, Exposure, NAAQS, $PM_{2.5}$, Vehicles*

1. Introduction

Air pollution has remained consistently high and rising in Delhi and has also shown a mixed trend over the years. During the initial years of the last decade, the Indian Supreme Court directives have helped arrest soaring air pollution trend in Delhi in terms of improvement of emissions standards, implementation of largest ever public transport strategy on CNG; capping on age of commercial vehicles; improvement in vehicle inspection survey and diversion of substantial truck traffic otherwise plying on the city roads. Following other remedial measures, Delhi government also relocated polluting industries, cut power plant pollution and banned open burning. Particulate matter (PM) is a complex mixture of solid and liquid particles that vary in size and composition, and remain suspended in the air. Over the past decade, many health effect studies have shown an association between exposure to PM and increases in daily mortality and symptoms of certain illnesses. In cities, PM is mainly a product of combustion from mobile sources such as cars, buses, trucks, and construction equipments, and from stationary sources such as heating furnaces, power plants, and factories (Gupta and kumar 2006). High levels of ambient air pollution have been associated with anthropogenic sources such as industry, rubbish burning, and traffic (Lestiani and Santoso, 2011; Santosa et al., 2008; Santoso et al., 2011)

Some PM is emitted directly into the atmosphere as particles (primary particles), while some particles are produced by chemical reactions in the atmosphere (secondary particles). The size of ambient air particles ranges over a wide scale, from approximately 0.005 to 100 µm in aerodynamic diameter (from the size of just a few atoms to about the thickness of a human hair). Particles fall into three basic size modes: ultrafine particles (smaller than about 0.1 µm in diameter), fine particles (between 0.1 and 2.5 µm), and coarse particles (larger than 2.5 µm). Transportation has been recognized as a significant source of ambient PM in urban areas (Charron and Harrison, 2005; Abu- Allaban et al., 2007). Air quality inside public transportation vehicles is an important issue for passengers because many commuters spend substantial amounts of time in these microenvironments. Levels of $PM_{2.5}$ inside different transportation vehicles have been widely studied (Adams et al., 2001). $PM_{2.5}$ is a subset of PM_{10} and includes those particles with an aerodynamic diameter less than 2.5 µm.

1.1 Health Impacts of Particulate Matter

Ambient air pollution has been identified as the fifth biggest cause of mortality in India (Lim et al., 2012). India registered an increase of about 12% in the number of deaths and about 3% in years of life lost between 2005 and 2010. The alarming death increase in India over the period from 2005 to 2010 is 72000 more deaths. The disease-wise percentage distribution of deaths attributable to ambient particulate matter pollution in India is depicted in Figure 1. Exposure to fine particulate matter is associated with an increase in allergic disorders, asthma, cognitive deficits, brain abnormalities, decreased lung function, cardiovascular disease, cardiopulmonary disease, and death in extreme cases (Brook et al., 2010; Brunekreef and Holgate, 2002; Dockery and Stone, 2007; Padhi and Padhy, 2008).

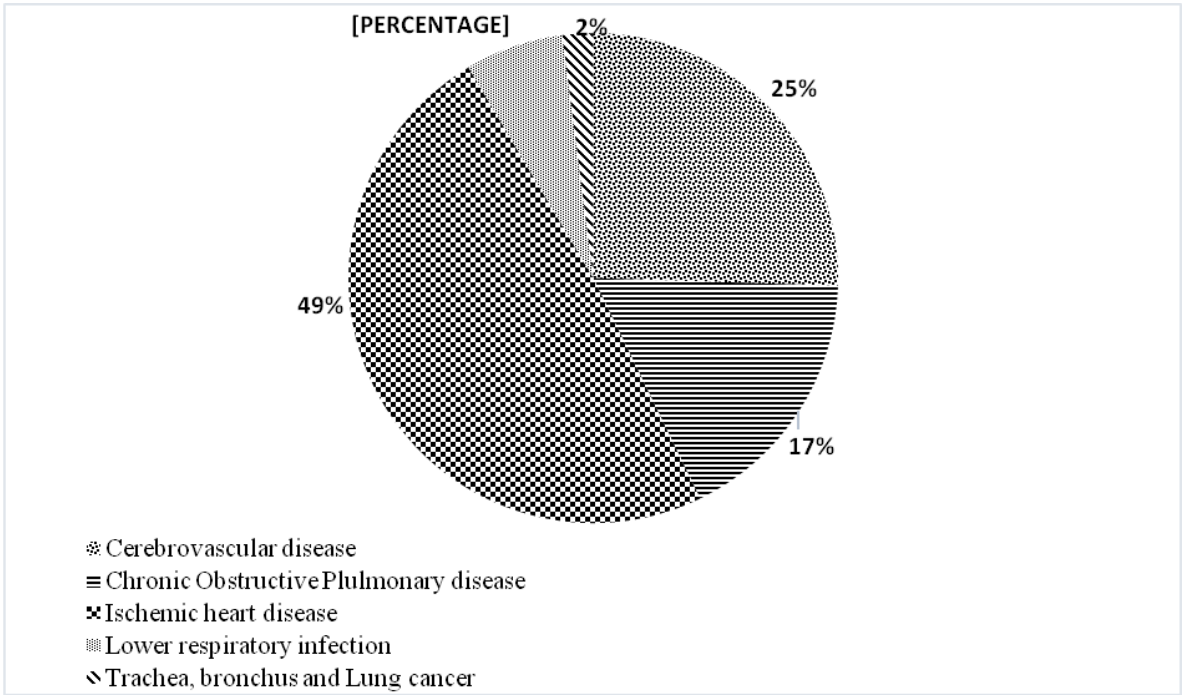


Fig.1. Percentage distribution of deaths from ambient PM pollution in India
Source: Lim et al, (2012)

As shown in Figure1, the highest value i.e., 49% of total deaths is attributable to Ischemic heart disease, while the lowest i.e., 2% deaths to trachea, bronchus and lung cancer. There have been many studies across the world and also in India establishing outdoor air pollution a serious environmental risk factor that causes or aggravates acute and chronic diseases. India shows that an increase in ambient air pollution significantly increases child morbidity (Gurjar et. al. 2008). The six cities considered in this study were Chennai, Delhi, Hyderabad, Indore, Kolkata and Nagpur. The study found that a rise in ambient air pollution significantly increases the likelihood of a child suffering from cough and fever. Controlling citywide air pollution could significantly lower child morbidity, and should receive greater emphasis in urban planning and infrastructure development. The adverse health effects of fine particulate matter ($PM_{2.5}$) have been identified by recent epidemiological studies (Han and Naeher 2006; Gascon et. al. 2016). Short and long-term exposure to $PM_{2.5}$ is associated with various health effects, including respiratory and cardiovascular diseases (Pope et al., 2002; Dominici et al., 2006). Additionally, exposure to $PM_{2.5}$ has been linked to induction of oxidative DNA damage in toxicological studies (Risom et al., 2005). Toxicological studies have also suggested that ultrafine particles (UFPs) may be more harmful to health than large particles because small particles have much larger adsorbed or condensed surface area than large particles with the same mass, and can deposit in alveoli, where they interact with epithelial cells (Delfino et al., 2005). Furthermore, fresh UFPs from traffic exhaust can damage DNA via systemic oxidative stress (Bräuner et al., 2007; Møller et al., 2008). Thus, exposure to ambient PM is a serious environmental risk for cardio-pulmonary and lung cancer mortality (Adams 2001b).

2. Study Methodology

The study domain is designated for various transport modes of commuting vehicles along a particular transport corridor. The method was based on a combination of studies conducted to assess the personal exposure of transport users in different cities of the world. The different transport modes chosen for the study were AC bus, Non- AC bus, AC car, Non- AC car and Auto-rickshaw. According to local authorities, these means of transportation represent major distribution in Delhi.

2.1 Selection of Route

The study was conducted along a particular transport corridor between Moolchand and Dr. Ambedkar Nagar in Delhi region, covering a distance of 5.8 Kilometre (Figure 2). The monitoring was conducted inside such transport modes/vehicles which are mostly used in day to day life by every commuter. The measurement of $PM_{2.5}$ was made during morning peak hours, evening peak hours and off-peak hours of the working days only.

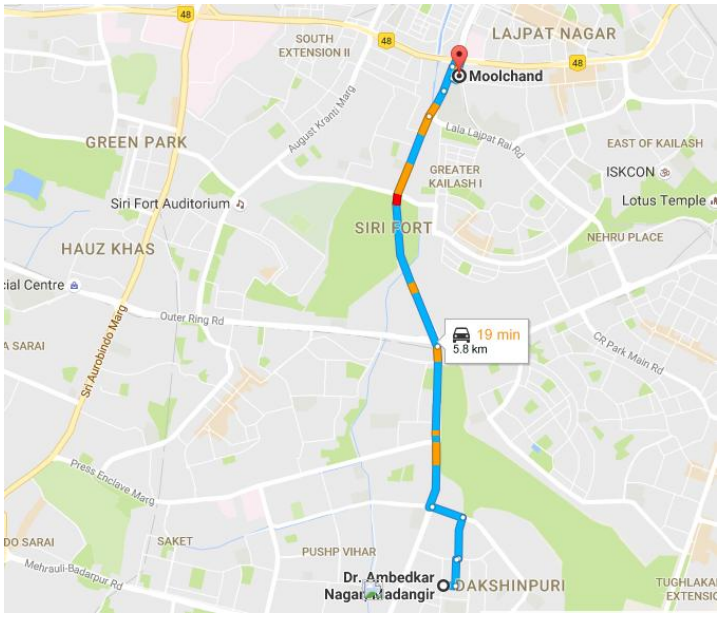


Fig. 2. Route map of study conducted in Delhi

2.1 Sampling and Measurement Equipment

Environmental Particulate Air Monitor (EPAM-5000) was used to monitor the $PM_{2.5}$ inside different modes/vehicles of transport along selected road stretch (Figure 3). The figure shows the overview of EPAM-5000 instrument. It is such kind of handy equipment that can directly measure the real time concentration of particulate matter. An affordable, easily deployable, and portable air monitor designed for measuring trace levels for ambient air pollutants. The unique sampling design allows for real-time data and filter gravimetric analysis utilizing the FRM 47mm Cassette located directly behind the optical sensor. EPAM-5000 offers a weather-tight carrying case, temperature-compensated electronics, easy to clean optical sensor and a standard 24-hour battery or continuous sampling with solar power panel accessory. $PM_{2.5}$ was sampled using this instrument at breathing level of commuters travelling on the selected modes of transport along identified route. The sampling of particulate matter was done 4 times in each mode of transport during morning peak hours, evening peak hours and off-peak hours of the different days.

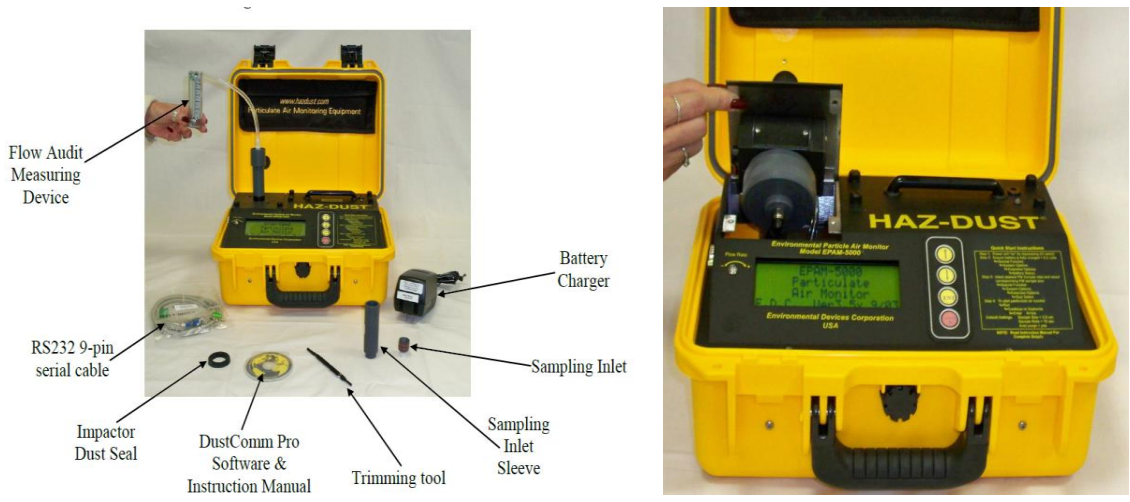


Fig.3. Haz-Dust™ Model EPAM-5000 systems

3. Result and Discussions

The average concentration of fine particulate matter in AC and Non-AC buses during different hours of the days is presented in Figure 4. During morning as well as evening peak hours of the days, the maximum concentration was found (93 µg/m^3) in Non-AC buses during 10:00 AM to 11:00 AM and 7:00 PM to 8:00 PM, whereas in AC buses, the maximum concentration of $PM_{2.5}$ was found as 56 µg/m^3 during 7:00 PM to 8:00 PM, quite lower than Non-AC buses. The figure shows the lowest concentration of 50.4 µg/m^3 in AC Buses during evening peak hour, whereas in Non-AC buses, the lowest concentration was found 62 % higher than the highest concentration of AC-Buses. The concentration measured in both types of buses during off-peak hours of the day was found slightly lower than the peak hours measured concentration. The concentration analysis in microenvironment of AC and Non-AC buses proves AC buses as a better alternative for the commuters who commute from one place to another place.

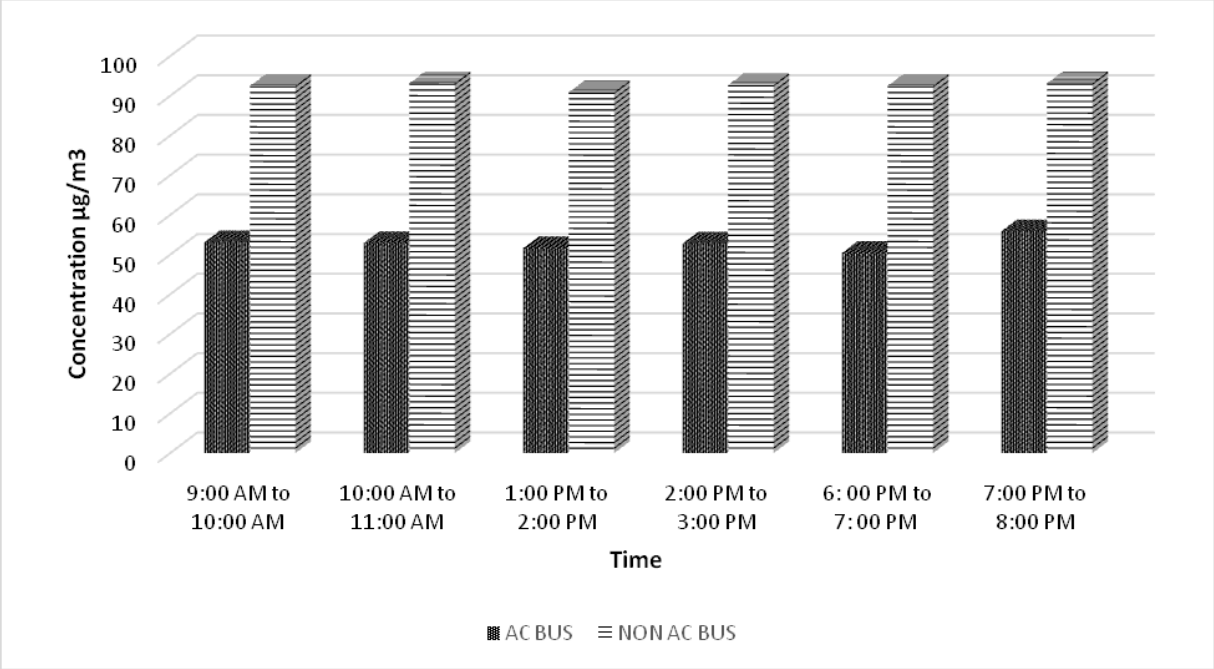


Fig. 4. Concentration of $PM_{2.5}$ in AC-Bus and NON- AC Bus during Morning, Evening and Off- Peak Hours

The concentration variation of $PM_{2.5}$ in AC and Non-AC Car is depicted in Figure 5. Without considering the model of the car, only concentration has been measured inside AC and Non-AC Car along the particular road stretch. Three modes of time interval i.e., morning peak hours, evening peak hours and off-peak hours of the days were chosen to conduct the study. The sampling was conducted at the back seat of the car at the height of breathing level. After analyzing the data of morning and evening peak hours along with off-peak hours, quite high difference of particulate matter concentration was observed inside AC and Non-AC car. From the figure, it is clear that the AC car is far better than Non-AC car in terms of exposure to particulate matter. The maximum average concentration inside AC car was found as 47.5 µg/m^3, whereas in Non-AC car, the highest concentration was 88.8 µg/m^3, which seems to be quite high than AC car. In addition to this, lowest concentration of $PM_{2.5}$ inside Non-AC car was nearly 50% higher than the concentration of particulate matter inside AC car. The average concentration range of particulate matter was recorded from 43.0 to 47.5µg/m^3 and 88.2 to 88.8 µg/m^3 in AC and Non-AC car respectively.

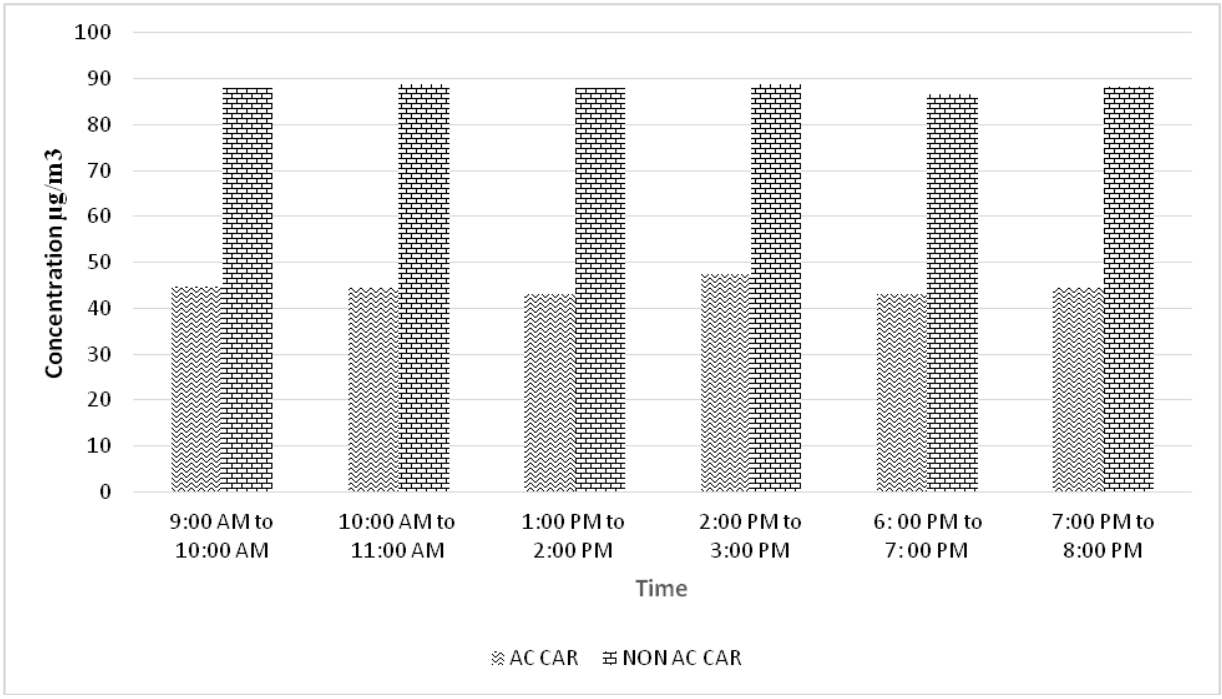


Fig. 5. Concentration of $PM_{2.5}$ in AC-Car and Non-Car in during Morning, Evening and Off- Peak Hours

The monitoring of $PM_{2.5}$ inside running auto-rickshaw along particular transport corridor during working days is presented in Figure 6. Since auto-rickshaw is a very common mode of travel in Delhi, due to that this mode was also considered to assess the exposure of commuters towards particulate matter. All three major divisions of time interval (morning peak hours, evening peak hours and off-peak hours) were chosen to conduct the study. After computing the sampled data of auto-rickshaw, quite high range of $PM_{2.5}$ concentration was observed. No major variation was observed amongst the average concentration of particulate matter during morning peak hours, evening peak hours as well as off-peak hours of the day. In comparison to AC bus, Non-AC bus, AC car and Non-AC car, 95.6 %, 10 %, 124 % and 14.8 % respectively, higher concentration of fine particulate matter was found inside auto-rickshaw. During morning and evening peak hours of the day, quite high concentration of $PM_{2.5}$ was observed inside auto-rickshaw, which may pose hazardous effects on human health of the commuters. During the entire study in auto-rickshaw, it was observed that the reported concentration of $PM_{2.5}$ always exceeded the National Ambient Air Quality Standards (NAAQS).

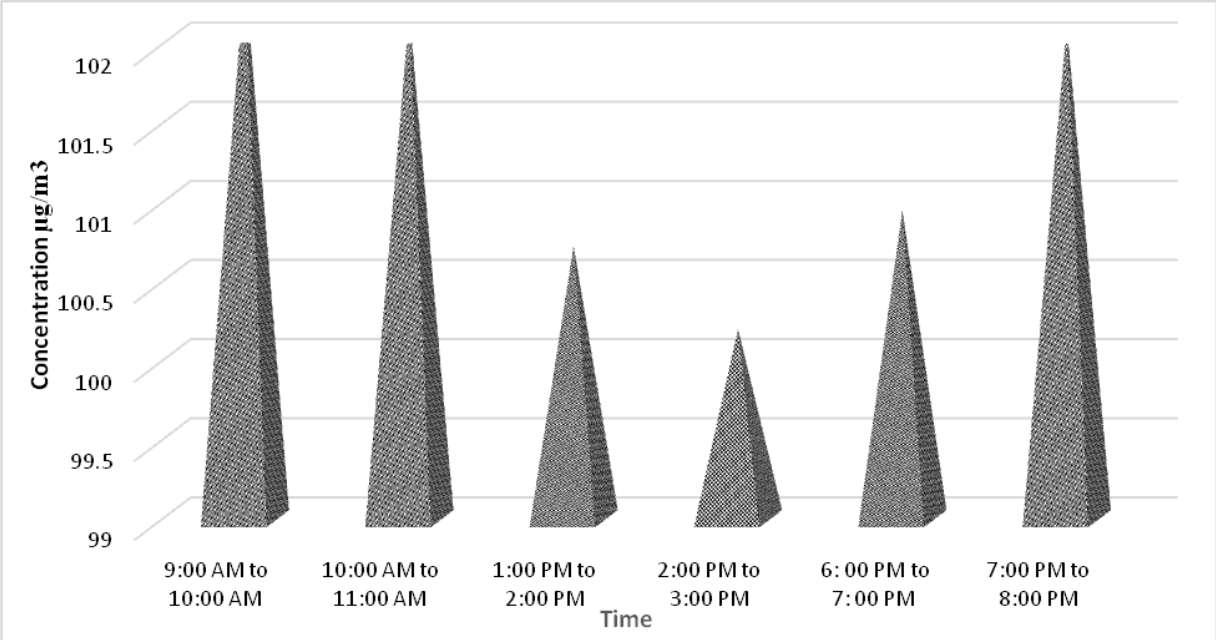


Fig. 6. Concentration of PM $_{2.5}$ in Auto during Morning, Evening and Off- Peak Hours

4. Conclusion

This study investigated commuters' exposure to $PM_{2.5}$, while commuting in different modes of transport like AC bus, Non-AC bus, AC car, Non-AC car and Auto-rickshaw. The measured concentration inside Non-AC bus as well Non-AC car was found much higher than the permissible limits. Among Non-AC vehicles, Non-AC bus reported average $PM_{2.5}$ as 92.8 µg/m^3 in morning as well as evening peak hours and 91.8 µg/m^3 in off-peak hours, while for Non-AC car $PM_{2.5}$ was found to be 88.5 µg/m^3 in morning peak hours, 88.5 µg/m^3 in off-peak hours and 87.5 µg/m^3 during evening peak hours and for auto as 102.3 µg/m^3 in morning peak hours, 101.5 µg/m^3 in evening peak hours and 100.5 µg/m^3 during off-peak hours.

On the other hand among the AC vehicles, slight variation was observed in the concentration of particulate matter between AC car and AC bus. AC car was found to be the most efficient among the rest with $PM_{2.5}$ detected as 44.6 µg/m^3 in morning peak hours, 43.8 µg/m^3 in evening peak hours and 45.3 µg/m^3 during off-peak hours followed by AC bus with average $PM_{2.5}$ detected as 53 µg/m^3 in morning, 53.2 µg/m^3 in evening and 52.2 µg/m^3 during off-peak hours. The reason behind this slightly higher concentration inside AC bus may be due to frequent opening and closing of doors, large size of engine, large surface area of bus, ridership of more people, frequent boarding and de-boarding of commuters etc.

Off-peak hour concentration of particulate matter was mostly reported less than those in morning and evening peak hours. Notably, among all commuting modes, auto was the worst with all $PM_{2.5}$ concentrations above 100 µg/m^3. The study revealed that mode type, location as well as temporal condition affect the exposure of commuters to $PM_{2.5.}$ It observed that close modes with air conditioning and segregation have reduced exposure level than open modes. Partially closed modes in mixed traffic lanes such as auto have higher $PM_{2.5}$ concentrations than others.

5. References

1. Abu-Allaban, M., Gillies, J.A., Gertler, A.W., Clayton, R., Proffitt, D., 2007. Motor vehicle contribution to ambient PM10 and PM2.5 at selected urban areas in the USA. Environmental Monitoring and Assessment 132, 155-163.
2. Adams, H.S., Nieuwenhuijsen, M.J., Colvile, R.N., 2001b.Determinants of fine particle (PM2.5) personal exposure levels in transport microenvironments, London, UK. Atmospheric Environment 35, 4557–4566.
3. Adams, H.S., Nieuwenhuijsen, M.J., Colvile, R.N., McMullen, M.A.S., Khandelwal, P., 2001. Fine particle (PM2.5) personal exposure levels in transport microenvironments London, UK. Science of the Total Environment 279, 29-44.
4. Bräuner et al., 2007. Exposure to ultrafine particles from ambient air and oxidative stress-induced DNA damage. Environmental Health, Perspectives 115, 1177-1182
5. Brook et al., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation, 121(21):2331–78.
6. Brunekreef, B., Holgate, S.T., 2002. Air pollution and health. Lancet, 360:1233–42.
7. Charron, A., and Harrison, R.M., 2005. Fine (PM2.5) and Coarse (PM2.5–10) Particulate Matter on a Heavily Trafficked London Highway: Sources and Processes. Environ. Sci. Technol. 39: 7768–7776.

8. Delfino, R.J., Sioutas, C., Malik, S., 2005. Potential role of ultrafine particles in associations between airborne particle mass and cardiovascular health. Environ Health Perspect, 113:934–46.
9. Dockery et al., 1993. An association between air pollution and mortality in six US cities. N Engl J Med, 329:1753–9.
10. Dockery, D., Stone, P., 2007. Cardiovascular risks from fine particulate air pollution. N Engl J Med, 356:511–3.
11. Dominici et al., 2006. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. Journal of the American Medical Association 295,1127-1134.
12. Gascon, M., Vrijheid, M., and Nieuwenhuijsen, M.J., 2016. The Built Environment and Child Health: An Overview of Current Evidence Curr Envir Health Rpt (2016) 3:250–257
13. Gupta, I., Kumar, R., 2006. Trends of particulate matter in four cities in India. Atmos Environ, 40:2552–66.
14. Gurjar, B.R., Butler, T.M., Lawrence, M.G., Levieveld, J., 2008. Evaluation of emissions and air quality in megacities. Atmos Environ, 42:1593–606.
15. Han, X., Naeher, L.P., 2006. A review of traffic-related air pollution exposure assessment studies in the developing world. Environ Int, 32:106–20.
16. Lestiani, D.D., Santoso, M., 2011. Analytical Methods INAA and PIXE applied to characterization of airborne particulate matter in Bandung, Indonesia. Atom Indones, 37:52–6.
17. Lim et al., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet., 380(9859):2224-60.
18. Møller et al., 2008. Air pollution, oxidative damage to DNA, and carcinogenesis. Cancer Letters 266, 84-97.
19. Padhi, B.K., Padhy, P.K., 2008. Assessment of intra-urban variability in outdoor air quality and its health risks. Inhal Toxicol, 20:973–9.
20. Pope, C., 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. J Am Med Assoc, 287:1132–41.
21. Risom, L., Møller, P., Loft, S., 2005. Oxidative stress-induced DNA damage by particulate air pollution. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 592, 119-137.
22. Santoso et al., 2011. Preliminary study of the sources of ambient air pollution in Serpong, Indonesia. Atmos Pollut Res, 2:190–6.
23. Santoso, M., Hopke, P.K., Hidayat, A., Dwiana, S.L., 2008. Sources identification of the atmospheric aerosol at urban and suburban sites in Indonesia by positive matrix factorization, 397 (1-3): 229-237.

.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

Comparison of Deflection and slope of Cantilever Beam with Analytical and Finite Element Method for Different Loading Conditions

Sanjay Kumar

Abstract— This paper presents the finite element formulation of cantilever beam and obtained deflection and slope and results are compared with Euler–Bernoulli beam theory (analytical solution). Here uniformly varying load (UVL) have been applied, deflection and slope are determined at different sections of beam. The analytical results are different with finite element method solutions for less elements but converges solution when number of elements are increased. The cantilever beam with uniformly distributed loading (UDL) is also analysed with both methods and results are compared to beam with UVL. The finite element method (FEM) and analytical solution are compared for different loadings. Here computer programme in matlab is used to obtain deflection and slope of Euler–Bernoulli beams at different locations for different loadings.

***Index Terms*— Cantilever, Euler–Bernoulli beam, uniformly varying loads, FEM, Matlab**

I. INTRODUCTION

The finite element method (FEM) is useful for many structural problem encountered during engineering application and applied science practical problem. But results obtained through FEM is not reliable like other numerical method. The accuracy of finite element method depends upon the discretization, interpolation function and number of elements. The finite element method consists three important steps: pre-processing, processing and post processing.[3]

In pre-processing, input data i.e. material properties, geometry, loads and boundary conditions are provided before simulation. Model is discretized into elements and meshes are generated automatically through matlab software. In processing stiffness matrix is generated and assembled to global matrix.

Euler-Bernoulli Beam Theory: The assumption considered for Euler- Beam theory is that plane section before and after bending remains plane and perpendicular to axis. The cantilever beam with uniformly varying load (UVL) is shown in the figure (1). Here loading q(x) is varying linearly q_o at fixed end to zero at free end on cantilever beam. At the free end of cantilever the point load (F_o) and moment (M_o) are acting in addition to UVL and UDL in figure (1) and figure (2).The differential equation of fourth order is used as governing equation for beam and transverse deflection is denoted by 'w'[1].

$$\frac{d^2}{dx^2}\left(EI\frac{d^2w}{dx^2}\right)=q(x)------(1)$$

Here, 'E' denotes the modulus of elasticity, 'I' is the area moment of inertia of beam about an axis perpendicular to both x and z axis and variation of loading $q(x)=q_o(1-x/L)$ where L is the total length of beam. The cantilever beam with uniformly distributed load (UDL) is shown in the figure (2)

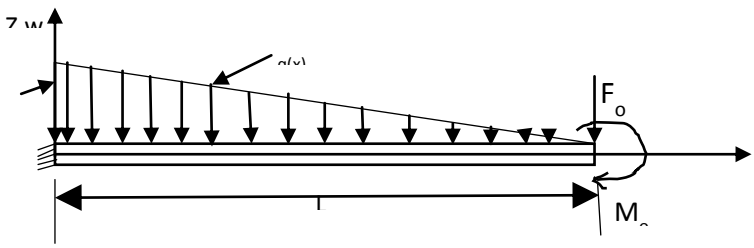


Fig (1) Cantilever beam with Varying distributed load

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

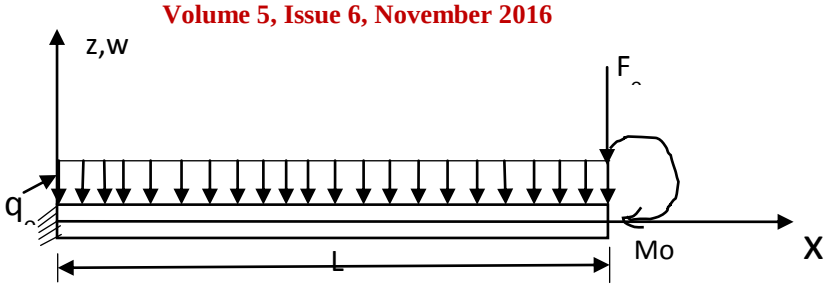


Fig (2) Cantilever beam with uniformly distributed load

The beam consists numbers of elements and considers an n^{th} element and distance of n^{th} nodes is at distance 'c' from global coordinates as shown in figure (3).

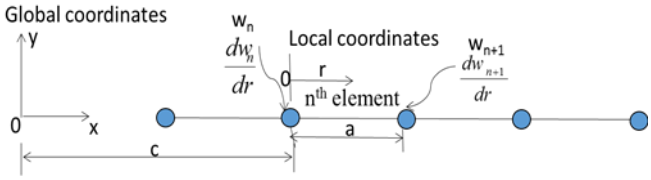


Fig (3) Degree of freedom for n^{th}element with local and global coordinates.

The transverse displacement (w) for element in z direction is written in terms of four constants c_1, c_2, c_3 and c_4

$$w = c_1 + c_2 r + c_3 r^2 + c_4 r^3 \text{ ---------- (2)}$$

Apply end conditions at both ends of element (figure (3)) and four shape functions/interpolations functions (N_1, N_2, N_3&N_4) are obtained as followings. These functions are also known as Hermite cubic interpolation functions.

$$N_1 = \left(1 - 3\frac{r^2}{a^2} + 2\frac{r^3}{a^3}\right),\quad N_2 = \left(r - 2\frac{r^2}{a} + \frac{r^3}{a^2}\right),\quad N_3 = \left(3\frac{r^2}{a^2} - 2\frac{r^3}{a^3}\right),\quad N_4 = \left(\frac{r^3}{a^2} - \frac{r^2}{a}\right) \text{--------(3)}$$

From equation (2), transverse displacement is expressed in terms of interpolation functions.

$$w = N_1 w_n + N_2 \frac{dw_n}{dx} + N_3 w_{n+1} + N_4 \frac{dw_{n+1}}{dx} \text{----(4)}$$

Use Galerkins approach/weighted residual methods to evaluate stiffness matrix and load vector. In this method integrate the weighted error over domain and equated to zero. Here weight function is considered same as 'w'.

$$\int_0^a (ERROR)(w)dr = 0 \text{--------(5)}$$

Element stiffness matrix (Ke) is obtained

$$Ke = \frac{EI}{a^3}\begin{bmatrix} 12 & 6a & -12 & 6a \\ 6a & 4a^2 & -6a & 2a^2 \\ -12 & -6a & 12 & -6a \\ 6a & 2a^2 & -6a & 4a^2 \end{bmatrix} \text{---(6)}$$

The elemental force matrix is obtained by integrating weighted load and obtained as

$$\{f^e\} = \int_0^a q_o \left(1 - \frac{r+c}{L}\right) w(r)dr$$

After integration final elemental load matrix becomes in form of column vector[1]

$$\{f^e\} = \frac{q_o a}{12}\begin{Bmatrix} 6 \\ a \\ 6 \\ -a \end{Bmatrix} - \frac{q_o a}{60L}\begin{Bmatrix} (9a+30c) \\ a(2a+5c) \\ (21a+30c) \\ -a(3a+5c) \end{Bmatrix} \text{---(7)}$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

Here, total number of elements is ten and number of nodes is eleven figure (4).The number within the circle represents the elements and simple numerals on black solid circle represents nodes. From above expression (7), it can be seen that load vector for UVL is not constant for every elements it changes from element to element. Elemental matrix is assembled for whole beam and it becomes global matrix [K]. Similarly elemental load matrix {F} is assembled and global load matrix is obtained.



Fig (4) Number of elements (10) and number of nodes (11)

For UDL the loading q_0 is constant throughout the length of beam, so elemental force matrix is obtained

$$\{f^e\} = \int_0^a q_o w(r)dr$$

$$\{f^e\} = \frac{q_o a}{12}\begin{Bmatrix} 6 \\ a \\ 6 \\ -a \end{Bmatrix} \text{--------(8)}$$

If global displacement is represented by $\{\delta\}$, then by Galerkin method, stiffness, displacement, distributed load and point load are related by relation (8)

$$[K]\{\delta\} = \{F\} + \{F_p\} \text{----(9)}$$

Here F_p is the external point loads acting on the nodes. Apply the boundary conditions at fixed node (node 1) and solution for displacements and rotations are obtained by use of equation (8). Matlab code has been prepared and used to obtain the solution.

The beam is made of steel materials having dimensions and loadings and mechanical properties are following:-length of beam is L=3m, Young modulus E=200x10^{6}kN/m^2, Moment of inertia about y axis, I=29x10^{-6}m^4,q_o=24kN/m,Point load at free end F_o=60kN,Moment at free end M_o=0kN-m [1].In diagram (2) length and mechanical properties of beam are same but loading conditions are different from figure(1),here uniformly distributed load(UDL) of 24 kN/m is

II. ANALYTICAL METHOD

The deflection of beam for UVL is obtained by direct double integration of moment at section "x" from fixed end as in equation (10). The slope of beam is obtained by direct integration of moment at section "x" from fixed end as in equation (11) [1]. The FEM and analytical solution for displacement and slopes are compared in the diagrams (5 to9).

$$w(x) = \frac{q_o L^4}{120EI}\left[10\left(\frac{x}{L}\right)^2 - 10\left(\frac{x}{L}\right)^3 + 5\left(\frac{x}{L}\right)^4 - \left(\frac{x}{L}\right)^5\right]$$

$$+ \frac{F_o L^3}{6EI}\left[-\left(\frac{x}{L}\right)^3 + 3\left(\frac{x}{L}\right)^2\right] + \frac{M_o L^2}{2EI}\left(\frac{x}{L}\right)^2 \text{--------(10)}$$

$$\theta(x) = -\frac{q_o L^3}{24EI}\left[4\left(\frac{x}{L}\right) - 6\left(\frac{x}{L}\right)^2 + 4\left(\frac{x}{L}\right)^3 - \left(\frac{x}{L}\right)^4\right] + \frac{F_0 L^2}{2EI}\left[\left(\frac{x}{L}\right)^2 - 2\left(\frac{x}{L}\right)\right] - \frac{M_0}{EI}\left(\frac{x}{L}\right) \text{----(11)}$$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

The deflection of beam for UDL is obtained by direct double integration of moment at section "x" from fixed end as in equation (12). The slope of beam is obtained by direct integration of moment at section "x" from fixed end as in equation(13) [2].

$$w(x)=\frac{q_o L^4}{24EI}\left[-6\left(\frac{x}{L}\right)^2+4\left(\frac{x}{L}\right)^3-5\left(\frac{x}{L}\right)^4\right]+\frac{F_o L^3}{6EI}\left[-\left(\frac{x}{L}\right)^3+3\left(\frac{x}{L}\right)^2\right]+\frac{M_o L^2}{2EI}\left(\frac{x}{L}\right)^2 \quad --(12)$$

$$\theta(x)=\frac{q_o L^3}{6EI}\left[-3\left(\frac{x}{L}\right)+3\left(\frac{x}{L}\right)^2-\left(\frac{x}{L}\right)^3\right]+\frac{F_o L^2}{2EI}\left[\left(\frac{x}{L}\right)^2-2\left(\frac{x}{L}\right)\right]-\frac{M_o}{EI}\left(\frac{x}{L}\right)-------(13)$$

The FEM and analytical solution for UVL loading for entire length of cantilever beam at step of 0.3m is shown in the table (1).Similarly for UDL loading displacement and slope is obtained and shown in table (2).

Table (1) Comparison of finite element method (FEM) solution with analytical for UVL.

Distance from fixed end X(m)	Displacement w(m)		Slope (Θ)	
	FEM	Analytical	FEM	Analytical
0	0	0	0	0
0.3	-0.0016	-0.0016	-0.0104	-0.0104
0.6	-0.0061	-0.0061	-0.0195	-0.0195
0.9	-0.0132	-0.0132	-0.0273	-0.0273
1.2	-0.0224	-0.0224	-0.0338	-0.0338
1.5	-0.0334	-0.0334	-0.0393	-0.0393
1.8	-0.0458	-0.0458	-0.0436	-0.0436
2.1	-0.0595	-0.0595	-0.047	-0.047
2.4	-0.0739	-0.0739	-0.0493	-0.0493
2.7	-0.0890	-0.089	-0.0507	-0.0507
3	-0.1043	-0.1043	-0.0512	-0.0512

Table (2) Comparison of finite element method (FEM) solution with analytical for UDL.

Distance from fixed end X(m)	Displacement w(m)		Slope (Θ)	
	FEM	Analytical	FEM	Analytical
0	0	0	0	0
0.3	-0.0016	-0.0016	-0.0104	-0.0104
0.6	-0.0061	-0.0061	-0.0195	-0.0195
0.9	-0.0132	-0.0132	-0.0273	-0.0273
1.2	-0.0224	-0.0224	-0.0338	-0.0338
1.5	-0.0334	-0.0334	-0.0393	-0.0393
1.8	-0.0458	-0.0458	-0.0436	-0.0436
2.1	-0.0595	-0.0595	-0.047	-0.047
2.4	-0.0739	-0.0739	-0.0493	-0.0493
2.7	-0.0890	-0.089	-0.0507	-0.0507
3	-0.1043	-0.1043	-0.0512	-0.0512

III. RESULT AND DISCUSSION

Figure (5) and (6) show the deflection and slope of beam respectively due to UVL from analytical and FEM method. The displacement and slope of cantilever beam obtained from analytically matched with finite element

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

solution for 10 elements due for UVL loading (figure (5)&(6)). The data obtained from analytical solution and FEM solution with help of MATLAB code [4]is mentioned in table (1) for UVL loading. First column indicates distance of nodes from fixed end to free of cantilever at step of 0.3m length. Second and third columns indicates value of displacement obtained from FEM and analytical respectively for UVL loading. The slope obtained from FEM and analytically is shown in fourth and fifth column in table(1).

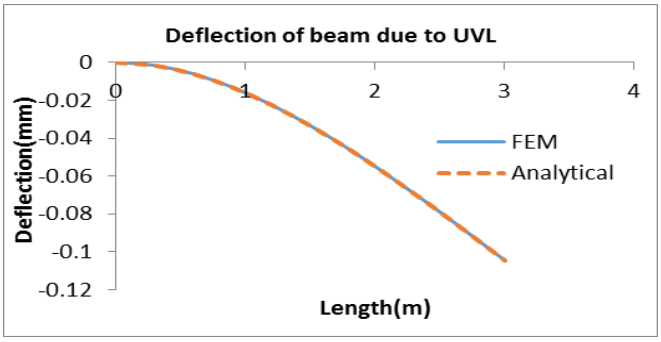


Fig (5): Comparison of deflection obtained analytically and through FEM for UVL.

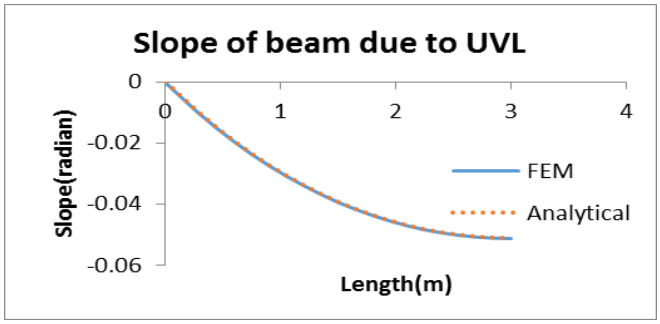


Fig (6): Comparison of slope obtained analytically and through FEM for UVL.

Figure (7) and (8) show the deflection and slope of beam respectively due to UDL from analytical and FEM method with use of MATLAB code [5]. The displacement and slope of cantilever beam obtained from analytically is slightly different with finite element solution for 10 elements due for UVL loading (figure (7)&(8)). This is always true for beam subjected to distributed loading is modeled by cubic displacement function. The solution obtained by beam theory is quartic (fourth order) polynomial but in FEM always it is assumes that displacement is cubical polynomial [2] .So finite element solution predicts stiffer structure than actual structure. The data obtained from analytical solution and FEM solution is mentioned in table (2) for UDL loading. First column indicates distance of nodes from fixed end to free of cantilever at step of 0.3m length. Second and third columns indicates value of displacement obtained from FEM and analytical respectively for UDL loading. The slope obtained from FEM and analytically is shown in fourth and fifth column in table(2).

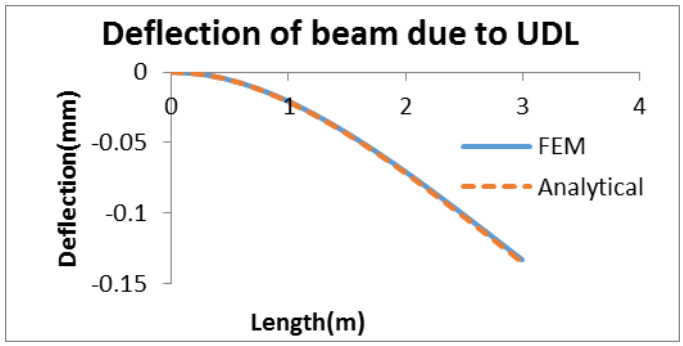


Fig (7): Comparison of deflection obtained analytically and through FEM for UDL.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

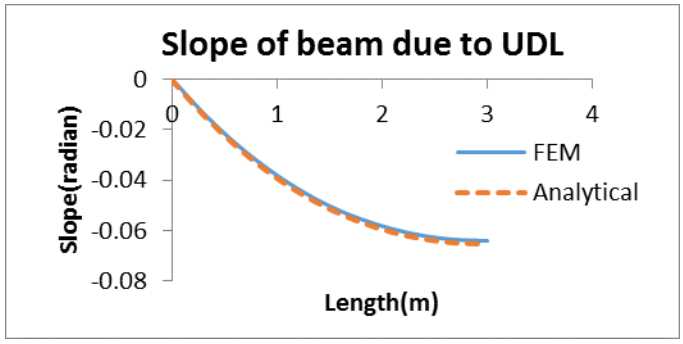


Fig (8): Comparison of slope obtained analytically and through FEM for UDL.

Figure (9) compare the deflection of beam due to UDL and UVL by FEM method for same loading. The displacement of cantilever beam due to UDL are more than UVL. The is true because loading in UDL is uniform thorough out the entire length of beam but in UDL it is maximum at the fixed end of cantilever and decreases to zero at the free end of the beam.

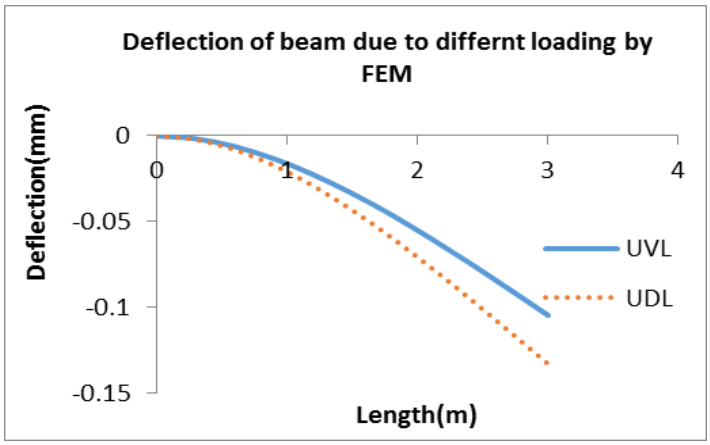


Fig (9): Comparison of deflection for different loadings.

Figure (10) shows slope of beam due to UDL and UVL by FEM method for same loading. The slope of cantilever beam due to UDL are more than UVL. The is true because loading in UDL is uniform thorough out the entire length of beam but in UDL it is maximum at the fixed end of cantilever and decreases to zero at the free end of the beam It is observed that deflection for UDL is greater than the UVL for same loading

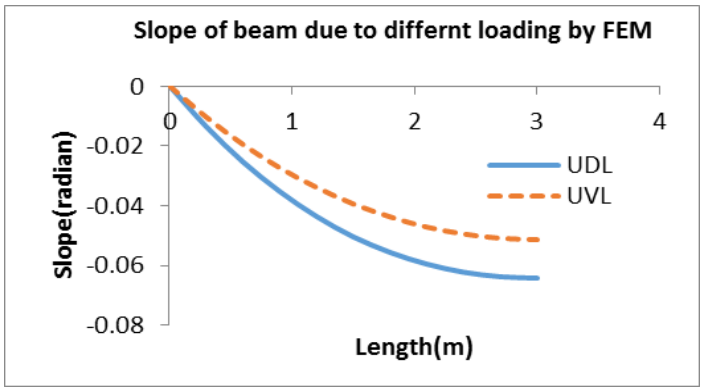


Fig (10): Comparison of slopes for different loadings.

IV. CONCLUSION

Following conclusions are obtained from this research.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 5, Issue 6, November 2016

- The solution for deflection and slope obtained from FEM and beam theory are exactly matched for ten elements with UVL loading
- The solution for deflection and slope obtained from FEM and beam theory are slightly lower for ten elements with UDL loading.
- Finite element solution predicts stiffer structure than actual structure.
- The deflection and slope are larger for UDL than UVL and increase from fixed end to free end for cantilever beam.

REFERENCES

[1] Reddy, J.N.(2013),An Introduction to the Finite Element Method , McGraw Hill, Education(India) Private Limited New Delhi.

[2] Logan,L Daryl (2010).A first course in the finite element method, Fourth edition. THOMSON.

[3] R.D.Cook,David.S.Malkus, Michel E Plesha, Robert J.Witt (2001).Concept and applications of Finite Element Analysis, Fourth edition.

[4] Kattan,I. Peter (2010) MATLAB Guide to Finite Elements :An Interactive Approach,Springer Internatinal Second Edition.

[5] A.J.M, Ferreira (2008) Matlab Codes for Finite Element Analysis: Solids and Structures, Springer.

AUTHOR BIOGRAPHY



Sanjay Kumar is working as Assistant Professor at Delhi Technological University, Delhi since September 2008. He has done B.Sc (Engg.) from M.I.T. Muzaffarpur in 1998 with Ist Div, M.Tech (Machine Design) from I.I.T. BHU, Varanasi in 2000 with Ist Div, and pursuing Ph.D from IIT Delhi. He is Life Member of ISTE. . He has published more than 10 Research Papers in various National/Internal Journal/ Symposium.. He has also worked as Assistant Professor in K.I.E.T, Ghaziabad for 8 years (2000-2008). He has guided 2 PG students and more than 30 UG students.

Volume 4, Issue 4(2016) 614-618 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Design of Solar Powered UAV

Janardan Prasad Kesari*, Abhishek Shakya
Department of Mechanical Engineering, Delhi Technological University, Delhi. India

Article Info
Article history:
Received 25 October2016
Received in revised form
20 November 2016
Accepted 28 November 2016
Available online 15 December 2016

Keywords:
Solar powered UAV, solar panel area, UAV, Design

Abstract
This paper summarizes the final project of undergraduate student of the faculty of Mechanical Engineering at the DTU, New Delhi, India. The project's aim is to design, build, test and fly a solar powered Unmanned Aerial Vehicle. Integrating solar energy into modern aircraft technology has been a topic of interest and has received a lot of attention from researchers over the last two decades. A few among the many potential applications of this technology are the possibility of continuous self-sustained flight for purposes such as information relay, surveillance and monitoring. The use of UAS is increasing rapidly due to the reduced production and operating cost compared to the large conventional aircraft.

Nomenclature

UAV – Unmanned Aerial Vehicle
MPPT – Maximum Power Point Tracker
AR – Aspect Ratio
AOA – Angle of Attack
C_l – Coefficient of lift
C_d – Coefficient of drag
C_m - Coefficient of momentum
b- Chord length
S- Wing surface area
P_{av}- Power consumed by avionics
$P_{electotal}$- Total electrical power needed by aircraft
P_{pld}- Power needed by payload
∂- Angle of declination
ω_s- Hour angle
H_0- Solar Irradiance
I_b- Hourly Total Radiation
T.O- Take Off
NACA- National Advisory Committee for Aeronautics
SD- Selig Donovan

1. Introduction

Possible applications of the Unmanned Aerial Vehicle (UAV) include military and classified surveillance flights communication links. Solar powered UAV can be employed in many of the above mentioned missions as it is capable of long endurance flight and does not require much maintenance. The Solar Powered UAVs use an unlimited power source for propulsion and other electrical systems. Using Photovoltaic (PV) cells, solar radiation is converted into electric power and then converted into kinetic energy by the electric motor. The main difficulty as for today is the low efficiency of both PV cells and motors. This paper presents the design of the Flare, a Solar Powered UAV where small aircrafts are difficult to be detected by radars. Scientific applications include ozone monitoring, and collection of data for weather and global warming studies. Commercial applications include aerial surveying, geological and topographical mapping.

2. Air Vehicle Design

2.1 Weight Estimation

The Gross Take-Off Weight was estimated and estimated weight for our solar powered UAV "FLARE" is 4kg.

***Corresponding Author,**
E-mail address: drjpkesari@gmail.com
All rights reserved: http://www.ijari.org

Table1. Weight estimation

Component	Weight (In Grams)
Wing	1410
Fuselage	330
Tail Boom	80
Tail Servos	78
Aileron Servos	70
Tail Servos	40
Autopilot	270
System Battery	360
Electric Motors	245
Speed Controller	40
Propeller + Spinner	20
PV cells	760
Wiring	200
TOTAL	3903 gm
Approx.	4kg

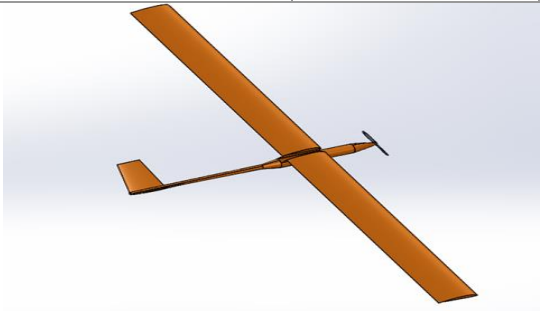


Fig.1 Solar powered UAV

2.2 Airfoil Selection

Airfoil Selection is one of the most crucial steps of aircraft designing. The desirable characteristics include high lift to drag ratio, low coefficient of moment and high stall angle. Different Low Reynolds number high lift airfoils were analysed in XFLR5 software and SD7032 was selected for high to drag ratio and high stall angle.

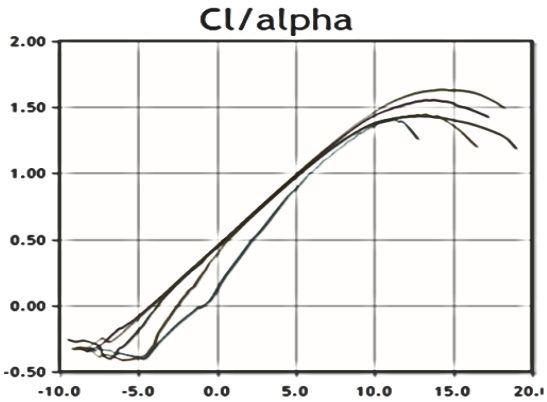


Fig.2 Coefficient of lift characteristics

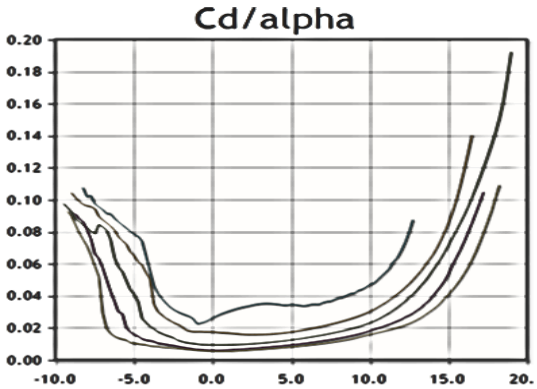


Fig.3 Coefficient of drag characteristics

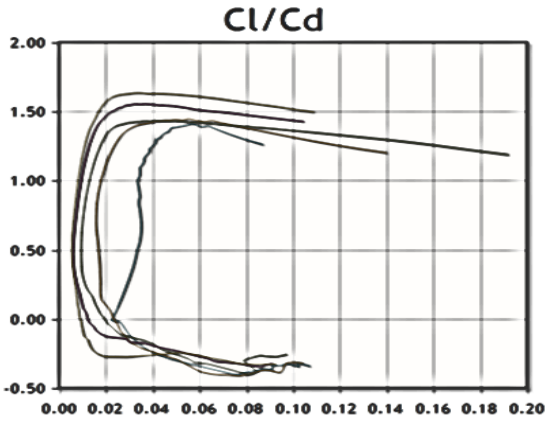


Fig.4 Coefficient of lift and drag characteristics

2.3 Sizing of the Plane

A GUI for calculating the dimensions of the UAV was made in MATLAB. The code takes airfoil parameters and design coefficients as its input and determines the dimensions of wing and tail based on standard sizing equations. The tail was calculated in a way to ensure that the derivative of coefficient of moment with respect to angle of attack is negative so that aircraft is longitudinally statically stable.

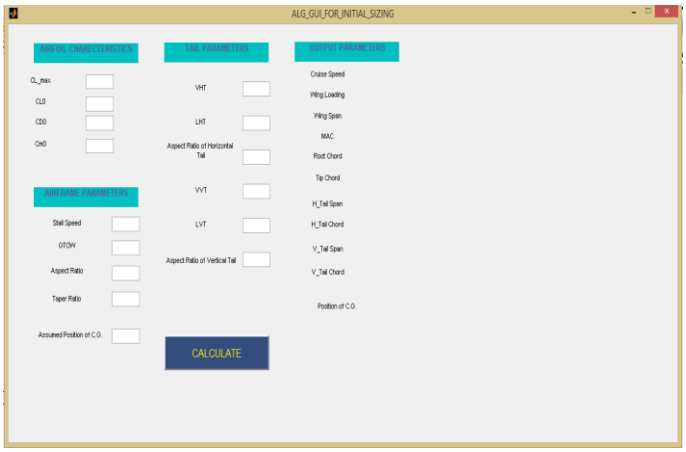


Fig.5 Spread sheet

Table 2: The final dimensions

Parameter	Value
Maximum Take-Off weight	4 Kg
Length	2.4m
Wing Airfoil	SD7032
Aspect Ratio	13.20
Wing Area	1.34m2
Wing Dihedral	3.5o
Tail Airfoil	Naca0006

3. Power Calculations

3.1 Estimation of Solar Panel Area

For the aircraft to be able to fly on any day of the year, it has to be designed on the coldest day of the year at a given location, which is defined by the latitude (/), and the common weather condition such as tropical, desert, or snow. The following procedure gives clear and straight forward methodology to estimate different parameters at given altitude and payload weight. The declination (d) is the angular position of the sun at solar noon with respect to the plane on the equator. The declination changes slightly each day and can be calculated using.

$$\partial = 23.45 \sin(360\frac{284+n}{365}) \quad (1)$$

The day of the year (n) has to be specified as part of a numeric sequence that starts with 1 for January 1st and 365 for December 31st. The hour angle of sunset (x) has to be estimated in order to evaluate the hours of sunlight for a given day. It is function of the latitude (/) and declination (d):

$$\omega_s = cos^{-1}(tan\emptyset * tan\partial) \quad (2)$$

The exact hour angle of sunset can be taken as the answer obtained from Eq. (2), whereas the sunrise can be taken as the negative of the same answer. The collected energy, at each hour during the day per unit area at a certain

location, changes dramatically with altitude. This is because near sea level there is a significant effect for the particulates and water vapor. On the other hand, at high altitudes, the cloud cover is negligible, thus there will be no daytime interruption of sun light.

For altitudes below 2.5 km, once the sunset hour angle has been found, the daily average total extraterrestrial irradiance ($\boldsymbol{H_0}$) availaible from the sun can be estimated using:

$$H_0 = \frac{24\times1367}{\pi}\left(1 + 0.033\, cos\frac{360n}{365}\right) * (cos\emptyset\, cos\partial\, sin\omega_s + \frac{\mu\omega_s}{180} sin\emptyset\, sin\partial) \quad (3)$$

The total hours of the day can be found using:

$$N = \frac{2}{15} cos^{-1}(-tan\emptyset\, tan\partial) \quad (4)$$

The average daily hours of bright sunshine (n) can be easily found from local weather stations. These data are based on Campell-Stokes instrument measurements. For this design, it was taken from India Meteorological Office At the coldest month of the year in Delhi, this value is 7.4. Then, the monthly average daily radiation on a horizontal surface can be estimated using:

$$\frac{H}{H_0} = a + b\frac{n}{N} \quad (5)$$

The above equation employs two empirical constants 'a' and 'b' which account for the local climate. A list of varies climates around the globe is given, from which the closest match was chosen for New Delhi. For New Delhi climate, a = 0.41, and b = 0.34. At each hour of the day (h), the instantaneous hour angle (x) can be estimated using

$$\omega = \frac{(15h-180)\pi}{180} \quad (6)$$

The hourly total radiation (I_b) per meter square at a certain hour, can be estimated using:

$$\frac{I_b}{H} = \frac{\pi}{24}(c + d\, cos\omega)\frac{cos\omega - cos\omega_s}{sinw_s - \frac{\pi\omega_s}{180}cos\omega_s} \quad (7)$$

The two constants in the above equation are given by:

$$c = 0.409 + 0.5016\sin(\omega_s - 60) \quad (8a)$$

$$d = 0.6609 - 0.4767\sin(\omega_s - 60) \quad (8b)$$

The radiation at each hour during the day has to be integrated to find the total energy from the sun in a day per meter squared.

Calculated: $I_b = 1181\frac{W}{m^2}$ (9a)

Experimental: $I_b = 1117\frac{W}{m^2}$ (9b)

3.2 Final Results:

Estimated solar panel area = $1.1m^2$

Efficiency of solar array = 15%

Power Generated by Solar array ≅ $180\ W$

3.3 Electrical characteristics

At irradiance of 1000 W/m^2

Table 3-Power Used

Open circuit voltage	0.670V
Short circuit voltage	5.9A
Maximum power voltage	0.560V
Maximum Power Current	5.54A
Rated power	3.1W
Efficiency	20%
Temperature Coefficients	
Voltage	-1.9$\frac{mV}{°C}$
Power	-0.38$\frac{\%}{°C}$

3.4 Propulsion Calculations:

3.4.1 Thrust and Power requirement

Using the aerodynamic calculations and assuming a 4kg vehicle weight required thrust and power were calculated and then translated to Motor Input Required Power using motor, gearbox and propeller efficiencies.Minimum required power for cruise is 60 [W] at 8[m/s]. Maximum cruise velocity requires 75W

3.4.2 Propeller Properties

The chosen propeller, a folder Aeronaut CAM, was selected using an electric propulsion system performance testing software, MotoCalc. After deciding on the belly landing concept a folding propeller was mandatory. Propeller diameters and pitches were checked for required thrust. The chosen propeller has a 15" diameter and 10" pitch.

3.4.3 Thrust required during level flight

Lift and Drag produced by vehicle at $C_{al=}1.082\ and\ C_{admax} = 0.052$ are summarized in table below:

$$Lift \quad : L = \frac{1}{2}\delta v^2 S C_{almax} \quad (8)$$

$$Drag \quad : D = \frac{1}{2}\delta v^2 S C_{admax} \quad (9)$$

Now, Thrust required by the aircraft is given by:

$$Thrust: \ T = \frac{W}{\frac{L}{D}}$$

$$= 1.885\ \text{N} \cong 2\ N \quad (10)$$

3.4.4 Power for Level Flight

Power required by the aircraft is given by:

$$v_{cruise} = \sqrt{\frac{\frac{2W}{S}}{\partial C_{al}}} \quad (11)$$

$$P = Tv \quad (12)$$

$$Thrust: \ T = \frac{W}{\frac{C_l}{C_d}} \quad (13)$$

$$P = \frac{W^{\frac{3}{2}}*\sqrt{2AR}*C_d}{C_l^{\frac{3}{2}}*\sqrt{\partial S}} \quad (14)$$

Hence the total energy can be calculated by taking into account the efficiencies of the components as follows:

$$P_{elec_{total}} = \frac{1}{\eta_{(ctrl+motor+propeller+gearbox)}} P_{level} + \frac{1}{\eta_{tr}}(P_{av} + P_{pld}) \quad (15)$$

$$P_{elec_{total}} = 88\ W$$

So, to fly the aircraft we need minimum 90W power per hour.

Thrust produced

$$T = \frac{\pi}{4}D^2\partial v\Delta v \quad (16)\ T = 14N$$

Propeller efficiency

$$\eta_{prop} = 65\%$$

$$T_{act} = 0.65 * T \ \ T_{act} = 8.8N$$

3.5 Power Plant and Solar Array

The values of parameters have been given in the table 4 and table 5:

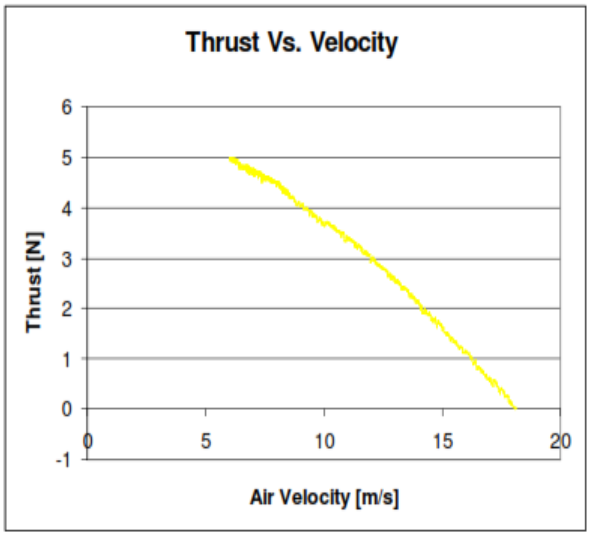


Fig .6 15"x10" Propeller. Thrust Vs. Velocity

Table 4: Power Plant

Parameter	Value
Electric Motor	Hacker B50-13S
Speed Controller	Hacker X-30
Gear Ratio	6.7:1
Propeller	15”X10”

Table 5: Solar Array

Parameter	Value
PV’s Area	1m2
PV’s Efficiency	20%
PV’s weight	0.76 Kg
PV’s max power	160 W

3.7 Aircraft’s Final Geometry

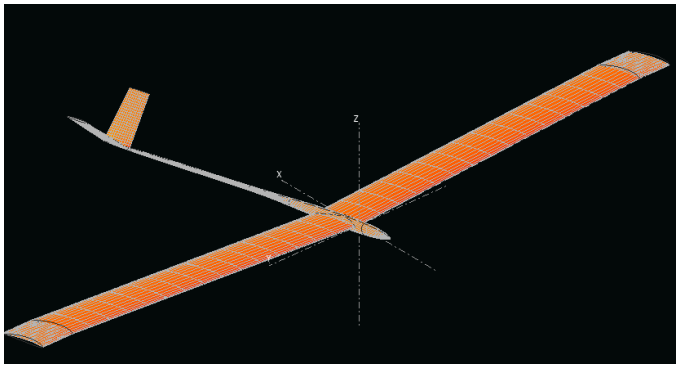


Fig. 7 Aircraft geometry

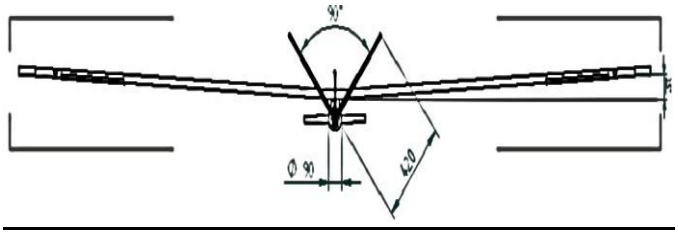


Fig. 8 Aircraft Angular Variation

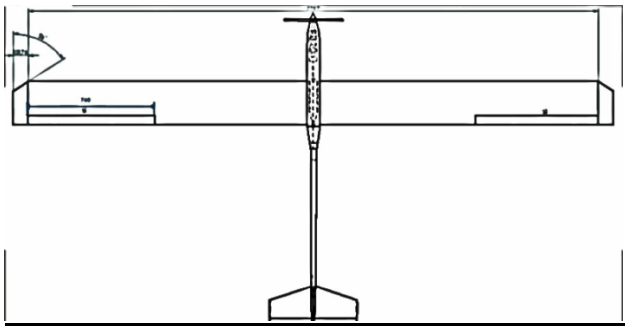


Fig. 9 Aircraft geometry for UAV

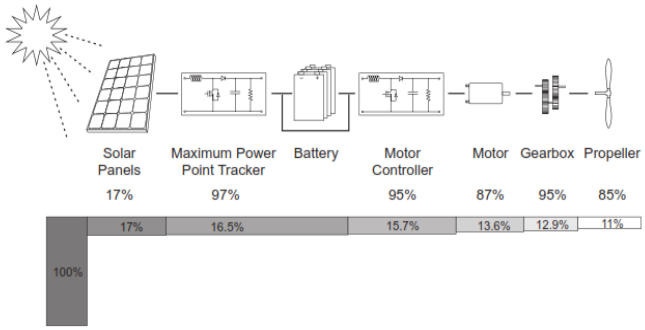


Fig. 9 Line diagram of actual working

3.8 Performance Parameters

Parameter	Value
Stall speed	6.6 m/s
Cruise Velocity	8 m/s
Take-Off Velocity	13 m/s
Power Required for Take-Off	80W
Max L/D	20.8
Aspect Ratio of Wing	13.2
Wing loading	28.98N/m2

4. Conclusions

This thesis presented a new methodology for the conceptual design of solar airplanes. It has the advantage to be very versatile and usable for a large range of dimension, from UAVs with less than one meter wingspan to manned airplanes. It is purely analytical and based on the concepts of energy and mass balances during one day using mathematical models that put the sizing of all elements on the airplane in relation. These models are used for efficiency or weight prediction and constitute a key part of such design method.

The methodology was used for the conceptual design of a prototype that would embed a small payload and with the

objective to prove the feasibility of continuous flight on Earth. It also allowed emphasizing some general principles. For example, it was clearly demonstrated that the most limiting technology at this time is the energy storage. Even with the best lithium-ion batteries, the energy storage constitutes more than 40 % of the airplane's gross weight. For that reason, what is critical for a continuous solar flight is not the day that has to be the longest, but the night that has to be the shortest.

References

[1] Kesari J P and P B Sharma and Lavnish Goyal, Transforming India Into a Green Nation: Bottom Up Approach through Solar Energy Education training and Innovation, Paper published in the proceedings of International Congress on Renewable Energy (ICORE) 2013 27-29 November , KIIT university , Bhubaneswar, Odisha.

[2]J. P Kesari and P.B. Sharma. Empowering Himalaya's Region with Solar Power Technologies, Education Research and Training. Proceedings of National Workshop on Renewable Energy Systems for Mountainous Regions: Issues & Challenges ReMount-2014, 14 Feb. 2014 Snow and Avalanche Study Establishment, DRDO, Chandigarh.

[3] J. P Kesari and P.B. Sharma. Empowering Himalaya's Region with Solar Power Technologies, Education Research and Training. Proceedings of National Workshop on Renewable Energy Systems for Mountainous Regions: Issues & Challenges ReMount-2014, 14 Feb. 2014 Snow and Avalanche Study Establishment, DRDO, Chandigarh.

[4] Raymer DP. Aircraft Design: A Conceptual Approach, 3rd edition, AIAAEducation Series.

[5] Etkin B. and Reid, LD. Dynamics of Flight (Stability and Control), JohnWiley & Sons, Inc, 1996

[6] J P Kesari 2015 Promoting Human Resources, Skill Development and Empowering India through Renewable Energy Education, Training and Innovation Keynote speaker and expert for the panel discussion. Proceedings of the conference on Make in India: Vision Strategy and Goal, 28-29 March, 2015 SDCOE

[7] Perkins, C.D. and Hage, R.E., Airplane Performance, Stability and Control, John Wiley & Sons, 1949

[8] N Baldock, MR Mokhtarzadeh-Dehghan. A Study of SolarPowered,High-Altitude Unmanned Aerial Vehicles. Aircraft Engineering and Aerospace Technology: An International Journal, 78(3), 2006,187–193.

[9] F Barbir, T Molter, D Luke. Regenerative Fuel Cells for Energy Storage : Efficiency and Weight Trade-offs. In Proc. International Energy Conversion Engineering Conference, AIAA-2003-5937, Portsmouth, VA, Auguste 17-21.

[10] Design of Solar Powered Airplanesfor Continuous FlightETH ZÜRICHfor the degree ofctor of Technical Sciencepresented byAndré NOTH

[11] P. Berry. The Sunriser - A Design Study in Solar Powered Flight. In Proc. of the World Aviation Conference, San Diego, USA, Oct 10-12 2000.

[12] SA Brandt, FT Gilliam. Design Analysis Methodology for SolarPowered Aircraft. Journal of Aircraft, 32(4):703–709, July-August 1995.

[13] JL Brocard, LNguyen. Image Processing on Mars Aerial Pictures for Sky-Sailor. Bachelor Thesis, Autonomous Systems Lab, EPFL, Lausanne, June 2005.

Ionics
DOI 10.1007/s11581-017-1984-2

ORIGINAL PAPER

Development and electrochemical performances of $Li_3V_2(PO_4)_3$ and $Li_4Ti_5O_{12}$ materials for lithium-ion battery

Rakesh Saroha [1] · Amrish K. Panwar [1] · Aditya Jain [1] · Jitendra Singh [1] · Sandeep Verma [1]

Received: 31 August 2016 / Revised: 12 January 2017 / Accepted: 15 January 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract The development and electrochemical performances of $Li_3V_2(PO_4)_3$ cathode and $Li_4Ti_5O_{12}$ anode materials have been reported in this work using CR2032 coin half-cell as well as full cell testing. Physicochemical characterizations of both the materials were studied using X-ray diffraction and scanning electron microscopy. The XRD patterns reveal the formation of the pure phase for both the samples. SEM micrographs show the formation of micrometer size, irregular-shaped and heterogeneously dispersed grains for $Li_3V_2(PO_4)_3$ (LVP) sample while $Li_4Ti_5O_{12}$ (LTO) possesses nanometer size grains with nearly spherical morphology. Electrochemical tests reveal that the full cell made with mass/capacity balancing factor $\gamma = 1.64$ delivers better rate and cycling performances as compared to the full cell with $\gamma = 0.92$. The initial discharge capacity was obtained to be 92 $mAhg^{-1}$ at 0.1 C for the cell with $\gamma = 1.64$ and retained 100% capacity even after 30 charge/discharge cycles at the same rate.

Keywords Cathode · Anode · Full cell testing · Rate performance

Introduction

Rechargeable lithium-ion batteries prove to be the backbone of the various consumer portable electronic devices and are now considered as a potential candidate in terms of power and energy density for automotive industry [1]. One of the key components of lithium-ion batteries is cathode material which plays a vital role in the determination of power and energy density. In order to develop enhanced cathode material with high energy density, various systems like layered transition metal oxide, spinel-type cathodes, and olivine-type cathodes have been studied extensively [2–4]. Among the above-stated cathode materials, $Li_3V_2(PO_4)_3$ (LVP) is one such material which possesses high theoretical capacity (197 $mAhg^{-1}$) when cycled between 3.0 and 4.8 V with an average working voltage of around 4.0 V [1, 5–10]. It contains three lithium sites from where three lithium ions can be completely extracted depending upon the V^{3+}/V^{4+} and V^{4+}/V^{5+} redox couples when cycled up to 4.8 V [9]. LVP exists in two different crystal structures, i.e., sodium superionic conductor (NASICON) or rhombohedral (space group: $R\bar{3}$) and monoclinic phase (space group: $P2_1/n$) [11, 12]. Monoclinic phase has been found to be thermodynamically more stable as compared to the NASICON phase and differs by the interconnection of "lantern" units [$V_2(PO_4)_3$] [13]. The NASICON phase of LVP consists PO_4 tetrahedra, and VO_6 octahedra joined together with their vertices to form lantern units [$V_2(PO_4)_3$] while three lithiums occupy three distinct crystallographic sites [14]. One lithium atom (L1) forms tetrahedral with four oxygen atom while L2 and L3 occupy distorted tetrahedral environment [15]. On the other hand, the monoclinic phase of LVP is quite similar to NASICON-type phase in crystallographic terms, but it is somewhat denser than NASICON phase. The three-dimensional structure in monoclinic phase is made from slightly distorted PO_4 tetrahedra and VO_6 octahedra by sharing oxygen vertices while lithium atoms are present at three sites in the interstitial voids [15]. Moreover, monoclinic LVP also exists in three different structural modifications depending upon the temperature viz. room temperature monoclinic α-phase, middle-

This paper has been presented at the 15th Asian Conference on Solid State Ionics, November 27–30, 2016, Patna, India

✉ Amrish K. Panwar
panwaramar@gmail.com

[1] Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

Published online: 23 January 2017

temperature monoclinic β-phase (400–460 K), and high-temperature orthorhombic γ-phase (460–573 K) [13].

Among available anode materials, spinel $Li_4Ti_5O_{12}$ (LTO) has proven to be a promising alternative anode material as it exhibits a flat charge/discharge plateau at a potential of approximately 1.55 V vs. Li/Li^+ [16–23]. In addition, it has a high theoretical capacity (175 $mAhg^{-1}$), structural stability, environmentally friendly, excellent cyclic stability, and low cost [24–27]. $Li_4Ti_5O_{12}$ has cubic spinel structure (space group: *Fd3m*) and Li^+ ions occupy *8a* tetrahedral sites while the *16d* octahedral sites were occupied by 5/6 Ti^{4+} and 1/6 Li^+ ions while O atoms held *32e* sites to form cubic close packing structure [28]. It shows a peculiar property of zero strain due to which it suffers from zero volume change in its unit cell during intercalation and deintercalation of lithium ions [29].

In this work, a systematic study of LTO/Li^+, LVP/Li^+ half-cells, and LTO/LVP full cell has been carried out. CR2032 coin cells were assembled for electrochemical tests which consist of LTO as anode and LVP as cathode materials. LVP and LTO electrode materials were synthesized using sol-gel and conventional solid-state methods, respectively. In the case of full cell, the mass/capacity ratio factor (γ) was determined using Eq. (1) and has been taken into consideration. It has been reported previously that for the optimized capacity value of a full cell, the two electrodes must be balanced in terms of mass or capacity (the so-called mass balancing) [30, 31]. The capacity or mass balancing condition can be expressed in terms of active mass ratios of the electrode materials as given by Eq. (1) [30]:

$$\gamma = \frac{m_+}{m_-} = \frac{\Delta x\, C_-}{\Delta y\, C_+} \tag{1}$$

where m_+ and m_- represent the active mass of the positive and negative electrodes, respectively, and the notations Δx and Δy represent the number of the lithium ions inserted into the anode and cathode, respectively, while C_- and C_+ represent the theoretical-specific discharge capacities of the negative and positive electrode materials, respectively.

Apart from the full cell, respective half-cells of the LTO and LVP electrodes were also assembled and tested for electrochemical performances. The physicochemical characterizations of the prepared samples were studied using XRD and SEM. Electrochemical tests were performed for all the assembled cells in order to compute specific discharge capacity and working voltage parameters.

Experimental

Material synthesis

Vanadium-based LVP cathode was prepared using the sol-gel route. Stoichiometric amounts of V_2O_5, CH_3COOLi, and $(NH_4)_2HPO_4$ (all of Sigma-Aldrich grade, ≥99%) were used as starting materials. Citric acid was used as a chelating agent while sucrose (GR MERCK) was used as a carbon source. Separate solutions of all the precursors were prepared in de-ionized water and mixed one by one in citric acid solution at 80 °C with continuous stirring at 500 rpm. The ratio of citric acid to metal ion concentration was kept at 1:1. The pH of the resulting solution was kept between 6 and 7 using ammonia solution. The amount of sucrose added was kept at 10 wt% of the metal ion concentration. After few hours, the resulting solution yields a transparent gel with mazarine type color. The resulting gel was then dried in vacuum oven at 80 °C for 12 h to obtain grayish floppy type xerogel. Thus, obtained material was then grounded and crushed using pestle and mortar and subsequently decomposed at 350 °C for 3 h under argon atmosphere in a tube furnace. The resulting powder was calcined at 700 °C for 15 h under inert flow which finally results into LVP powder.

For the synthesis of spinel LTO, titanium dioxide (TiO_2) and lithium carbonate (Li_2CO_3) were used as starting materials. LTO powder was prepared using the conventional solid-state method. The precursors were mixed using pestle and mortar for 7 h with acetone as a grinding medium. The molar ratio of titanium and lithium precursor's metal were kept at 5:4. Sucrose was used as a carbon source and the amount was kept at 10 wt% of the metal ion concentration. The resulting powder was subjected to two-stage thermal treatment. In the first stage, the powder was decomposed at 500 °C for 6 h in air. The powder thus obtained was grounded and subjected to annealing at 800 °C for 20 h. This results into white LTO powder.

Material physicochemical characterization

The crystal structure and phase confirmation of the as-synthesized samples were studied using RIGAKU make ULTIMA IV X-ray diffractometer (XRD) employing $CuK_{\alpha 1}$ radiation (λ = 1.540 Å). The XRD data was collected over the angular range of 10–70° with a step size of 0.02°. The microstructural properties of the prepared samples were studied using HITACHI make S-3700 N scanning electron microscope operated at 15 kV.

Electrode preparation

For electrochemical tests, CR2032 coin half-cells and full cells were assembled inside an argon-filled glove box workstation. The working LTO electrodes were prepared by dispersing 70 wt.% active material, 15 wt.% conductive carbon (super P), and 15 wt.% polyvinylidene difluoride (PVDF as a binder) in *N*-methyl-2-pyrrolidinone (NM2P) as a solvent. The obtained slurry was stirred for 5 h and then evenly coated on Cu-foil current collector using the "doctor blade" technique. For LVP electrodes,

the resulting mixture was prepared by dissolving 90 wt.% active material, 5 wt.% super P as a conductive carbon, and 5 wt.% PVDF in NM2P for 5 h and then coated on the Al-foil current collector. Thus, prepared coatings were dried in a vacuum oven at 120 °C for overnight to vaporize the solvent and then coatings were pressed using roller pressing machine at 50 MPa. Then circular electrodes with diameter 16 mm (for both cathode and anode) were punched and weighed. The coating thickness of the LVP cathode was modified by adjusting the thickness of doctor's blade during the slurry coating process in order to have requisite mass loading and balancing condition given by Eq. (1). The thickness of the active LTO anode electrode was found to be 55 ± 2 μm which results in an active mass loading of 2.5 ± 0.2 mg. The thickness of the active LVP cathode for the two coatings was determined to be 61 ± 2 and 85 ± 3 μm which gives an active mass loading of around 2.3 ± 0.1 and 4.1 ± 0.3 mg.

Coin cell assembly

The electrodes thus obtained were dried under vacuum at 80 °C to remove vapor content and finally brought inside the Mbraun make glove box workstation with O_2 and H_2O level ≤0.5 ppm. The CR2032 coin cells were assembled for electrochemical studies of the prepared samples. Half-cells corresponding to LTO/Li^+ and LVP/Li^+ were assembled and abbreviated as LTOH and LVPH ("H" stands for half-cell). Metallic lithium (Li/Li^+) was used as both counter and the reference electrode, and a microporous polypropylene sheet (Celgard 2400) was used as a separator. Two full CR2032 coin cells were also assembled with LVP as positive electrode and LTO as negative electrode. The two assembled cells were abbreviated as LTO/LVP1F and LTO/LVP2F ("F" stands for the full cell). The LTO/LVP1F cell consists of LTO anode having 2.5 ± 0.2 mg mass loading and LVP cathode with 2.3 ± 0.1 mg active mass. Likewise, LTO/LVP2F cell contains LVP cathode with 4.1 ± 0.3 mg active mass while anode remains the same. The electrolyte used was 1 M $LiPF_6$ in EC:DMC 1:1 *v/v*.

Electrochemical characterization

Cyclic voltammetry (CV) experiments were carried out over the potential range of 1.3–2.9 V (vs. Li/Li^+ metal) for the two full coin cells at a scan rate of 0.05 mVs^{-1}. The potential window for LTO half-cell was kept from 0.6–2.5 V vs. Li/Li^+ while the potential range was 2.8–4.8 V for LVP half-cell and data was collected at a scan rate of 0.05 mVs^{-1}. The galvanostatic charge/discharge (GCD) tests were carried out for all assembled cells. The theoretical capacity for LTO anode and LVP cathode has been taken as 175 and 197 $mAhg^{-1}$, respectively [1, 19]. The maximum electron uptake is 3 for both LVP and LTO samples while calculating theoretical-specific capacity [13, 32]. All the electrochemical tests were performed using Biologic Science Instruments make VMP3 model at 298 K.

Results and discussion

Structural and morphological analysis

The XRD patterns of the as-synthesized samples are shown in Fig. 1. From XRD patterns, it is found that all the diffraction peaks for LTO sample belong to cubic symmetry with *Fd3m*

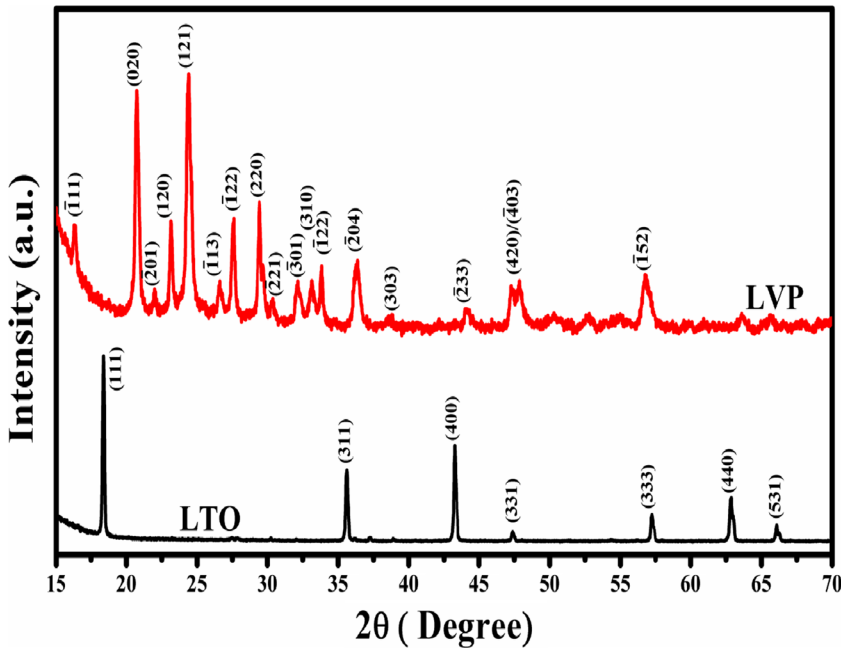


Fig. 1 XRD patterns of synthesized LVP/C and LTO samples

Table 1 Lattice parameter values for as-synthesized samples. All distance and lattice parameter units are in Å (10^{-10} m)

Sample	a (Å)	b (Å)	c (Å)	V $(Å)^3$
LTO	8.365	8.365	8.365	585.326
LVP	8.590	8.542	11.876	874.471

space group while diffraction peaks for LVP cathode are indexed to monoclinic phase with $P2_1/n$ space group. The sharp and intense peaks of the synthesized samples reveal that samples are highly crystalline in nature. The lattice parameter values for LVP cathode and LTO anode samples were obtained using X'pert high score plus software and summarized in Table 1. From the table, it can be seen that LTO belongs to cubic symmetry as $a = b = c = 8.365$ Å while LVP indexed to monoclinic phase with $a = 8.590$ Å, $b = 8.542$ Å, and $c = 11.876$ Å.

Figure 2a, b shows the obtained SEM micrographs for the LVP cathode and LTO anode samples, respectively. It can be observed from the SEM micrographs that the grains in LVP cathode (Fig. 2a) have an irregular shape and are sub-micron sized. The average grain sizes for LVP and LTO samples were obtained using ImageJ software and found to be 0.3–0.4 μm and 70–80 nm, respectively. However, for LTO (Fig. 2b) sample, the grain morphology is nearly spherical and nanometer size grains were obtained. The average crystallite sizes for the LVP cathode and LTO anode samples were also calculated using Scherrer's formula $d = k\lambda/\beta \cos\theta$, where d is the mean crystallite size, k is the constant parameter of shape factor (0.94), λ is the wavelength of X-ray (0.154 nm), β is the full-width half maximum, and θ is the corresponding Bragg diffraction angle. The average crystallite sizes for the cathode and anode were calculated to be 110 and 55 nm, respectively.

Electrochemical analysis of LVP and LTO half CR2032 coin cell

The active cathode and anode materials were tested with half-cell as well as full cell configuration in order to analyze electrochemical performances. Figure 3a, b displays the CV and first charge/discharge curves of the LVP and LTO active materials vs. Li/Li^+. LTOH anode material with an active mass loading of 2.5 ± 0.2 mg was scanned at 0.05 mVs^{-1} in the voltage window of 0.6–2.5 V while CV data was collected over the voltage range of 2.8–4.8 V for the active LVP cathode half-cell with 2.3 ± 0.1 mg active mass. It can be seen from the CV curves of LVP sample (inset of Fig. 3a) that charging (intercalation) process consists of four sharp peaks at 3.62, 3.7, 4.1, and 4.56 V which corresponds to the removal of three lithium ions from the LVP structure according to the reactions described somewhere else [13]. The first lithium ion was removed in two steps as depicted by two peaks at 3.62 and 3.70 V. The second and third lithium ions are extracted at 4.11 and 4.56 V, respectively. During the reverse process, three peaks around 3.50, 3.58, and 3.94 V were observed. The voltage peaks at 3.50 and 3.58 V during discharging represents the insertion of the first lithium ion while the peak at 3.94 V shows the insertion of the second lithium ion. However, the peak corresponding to the insertion of the third lithium ion from $V_2(PO_4)_3$ to $LiV_2(PO_4)_3$ phase is absent in the CV curves. It clearly indicates that the redox reactions involved during intercalation/deintercalation of the lithium ion are not symmetric. Moreover, the intensity of the peak around 4.1 V was found to be highest among all the peaks which show that the lithium-ion intercalation at this voltage will be maximum. The CV profile of LTO anode half-cell, i.e., LTOH consists of sharp anodic and cathodic peaks around 1.68 and 1.47 V, respectively. The symmetric nature of anodic and cathodic peaks for LTO anode shows that the redox reactions are reversible.

Figure 3b shows the first galvanostatic charge/discharge (GCD) curves of as-synthesized LVP and LTO half-cells vs. Li/Li^+. The voltage window for the LTO sample was kept 1.0–3.0 V while the voltage range was 2.5–4.8 V at 0.1 C rate for the LVP half-cell, respectively. It can be observed from Fig. 3b that the first charge curve of the LVP sample consists of three plateaus around 3.6, 4.1, and 4.6 V. The length of the flat plateau at 4.1 V is more as compared to the plateau at 3.6 V while the plateau at 4.6 V was found to have a little slope. These results indicate that the lithium-ion intercalation at these

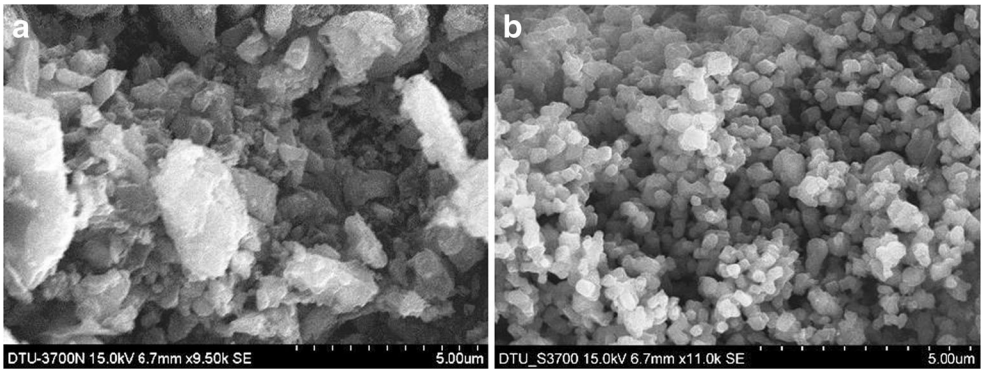


Fig. 2 SEM micrographs of LVP/C and LTO samples. **a** LVP/C. **b** LTO

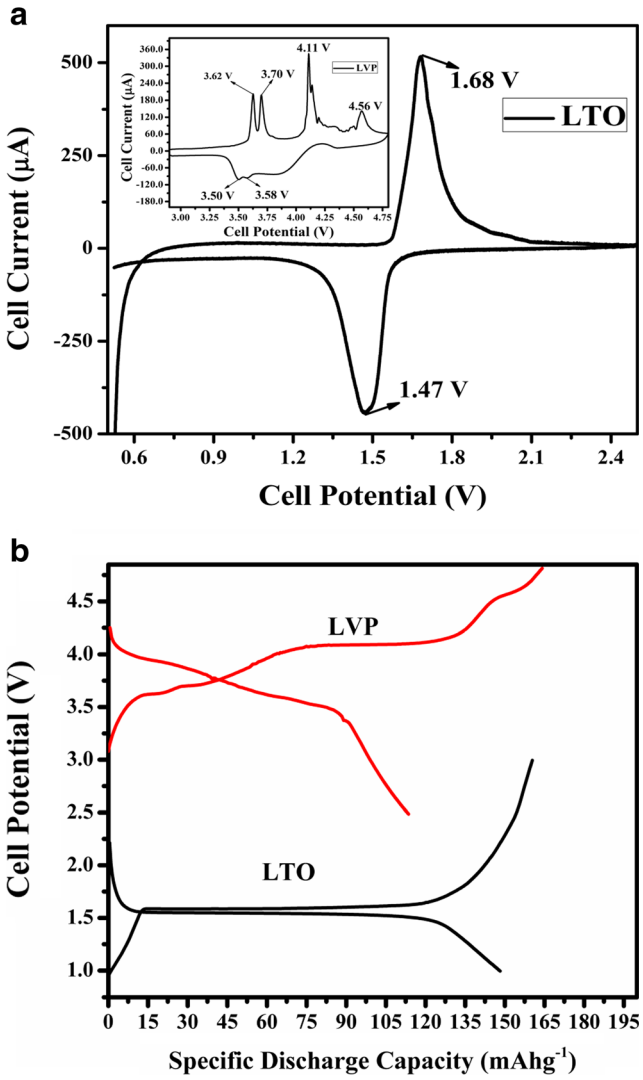


Fig. 3 **a** CV profiles at a scan rate of 0.05 mVs^{-1} and **b** initial charge/discharge at 0.1 C rate of LVP and LTO vs. Li/Li^+ half-cells

three-voltage plateaus follows different mechanisms. The three-voltage plateaus clearly suggest that every time when a lithium ion moves from cathode to anode during charging, the crystal structure of the cathode material undergoes modifications as oxidation state of vanadium changes. It can be seen from the first discharge of LVP sample that during discharging, only one plateau with significant slope appears which spans a voltage range of 4.2–3.1 V. It clearly shows that the redox reactions taking place inside the host are not exactly symmetric and reversible as predicted by CV results. The discharge capacity for LVP was observed about 114 (±2) mAh/g (~58% of the theoretical capacity) at 0.1 C rate. Likewise, the discharge capacity of the LTOH anode material was observed to be 150 (±5) mAh/g (~86% of the theoretical capacity) at 0.1 C rate.

Rate and cycling performances of the assembled half cells are shown in Fig. 4a, b, respectively. From Fig. 4a, it can be observed that LTO and LVP half-cells deliver a discharge capacity of 150 (±5) and 114 (±2) mAh/g, respectively, at 0.1 C rate. Even at a high discharge rate of 2 C, the LTO and LVP half-cells deliver nearly 47 and 51% of discharge capacity observed at 0.1 C rate. Figure 4b shows the cycling performance of the two half-cells at 0.1 C rate for 30 charge/discharge cycles. It can be seen that both LTO and LVP half-cells deliver nearly 97% of discharge capacity even at the 30th charge/discharge cycle. These results clearly demonstrate the integrity of the anode and cathode material.

Electrochemical analysis of LVP and LTO full coin cell

Two full CR2032 coin cells were fabricated with LTO as an active anode while LVP as cathode anode. The LVP to LTO mass balance was kept at 0.88:1.0 to test the full lithium-ion

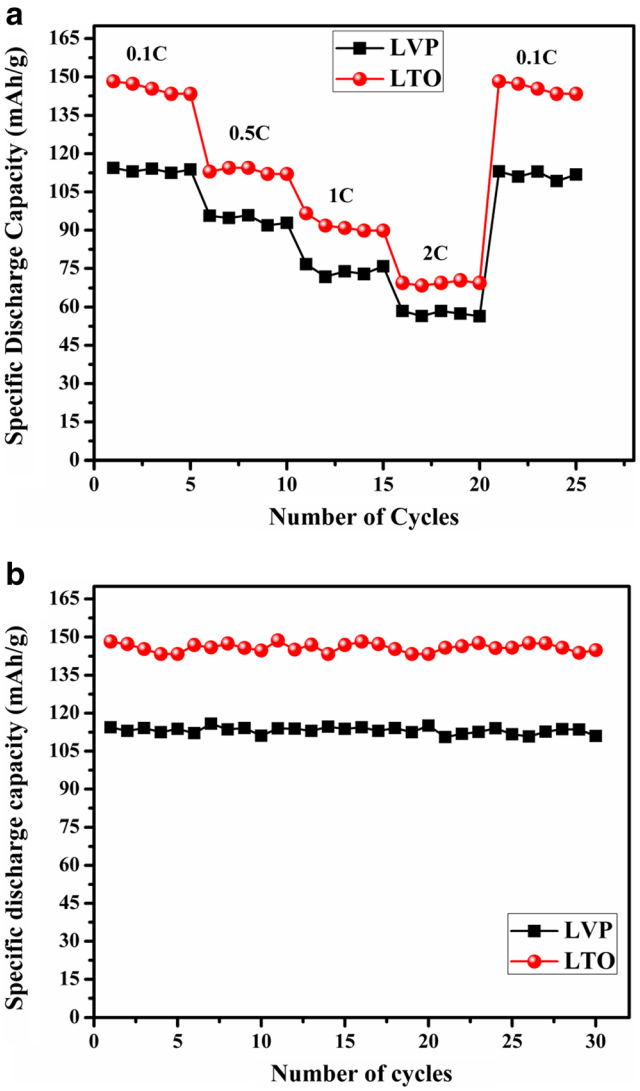


Fig. 4 **a** Rate performance of LVP and LTO half-cells at different C rates. **b** Cycling performance of LVP and LTO vs. Li/Li^+ half-cells at 0.1 C rate for 30 charge/discharge cycles

battery. However, the thickness of the LTO anode electrode was kept constant while the cathode electrode thickness was varied as 61 ± 2 and 85 ± 3 μm which gives an actual mass loading of around 2.3 ± 0.1 and 4.1 ± 0.3 mg, respectively. The mass/capacity ratio (γ) was calculated to be 0.92 for LTO anode and 61 ± 2-μm-thick cathode (abbreviated as LTO/LVP1F) while γ was found to be 1.64 for same LTO electrode and 85 ± 3-μm-thick cathode (abbreviated as LTO/LVP2F cell). Figure 5a shows the CV curves of LTO/LVP1F and LTO/LVP2F cells at a scan rate of 0.05 mVs^{-1} in the voltage window of 1.3–2.9 V. The redox peaks in the CV profile of the full cell mainly represent the LVP cathode. However, it can be seen that among two full cells, LTO/LVP2F possesses more intense cathodic/anodic peaks as compared to the LTO/LVP1F cell. From CV curves, it can be observed that LTO/LVP1F cell has one intense peak around 2.25 V and weak broad peak at 2.58 V during charging while during discharging, peaks were located at around 2.34 and 1.76 V. However, for LTO/LVP2F cell, one shoulder peak at 2.16 V along with peak at 2.25 V has also appeared during charging. The electrochemical reactions which are taking place inside LVP/LTO full battery can be described as [33]:

$$Li_3V_2(PO_4)_3 \rightarrow LiV_2(PO_4)_3 + 2Li^+ + 2e^- \quad (2)$$

$$\frac{2}{3}Li_4Ti_5O_{12} + 2Li^+ + 2e^- \rightarrow \frac{2}{3}Li_7Ti_5O_{12} \quad (3)$$

However, the reduced intensity of peaks during discharging shows that lithium deintercalation is not similar to lithium intercalation during charging which is also evident from the asymmetric nature of anodic and cathodic peaks. Moreover, the nearly superimposing CV profile of the two full cells reveals the decent reversibility of the redox reactions. The first charge/discharge curves of the prepared full cells at 0.1 C rate are shown in Fig. 5b. LTO/LVP2F shows better charge/discharge capacity profile as compared to LTO/LVP1F as

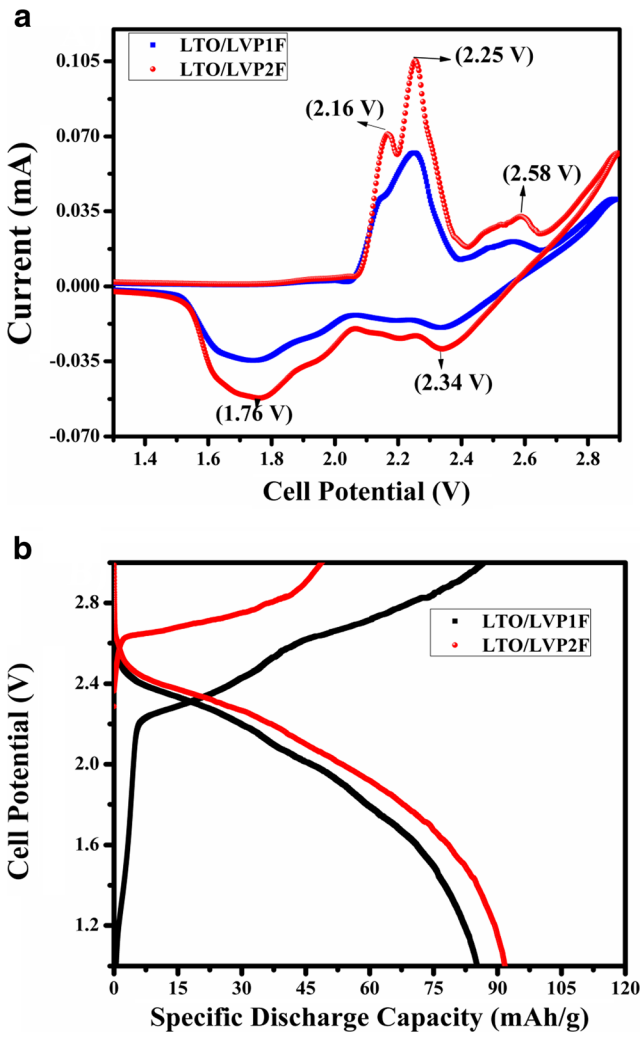


Fig. 5 **a** CV profiles at a scan rate of 0.05 mVs^{-1} and **b** initial charge/discharge at 0.1 C rate of LTO/LVP1F and LTO/LVP2F cells

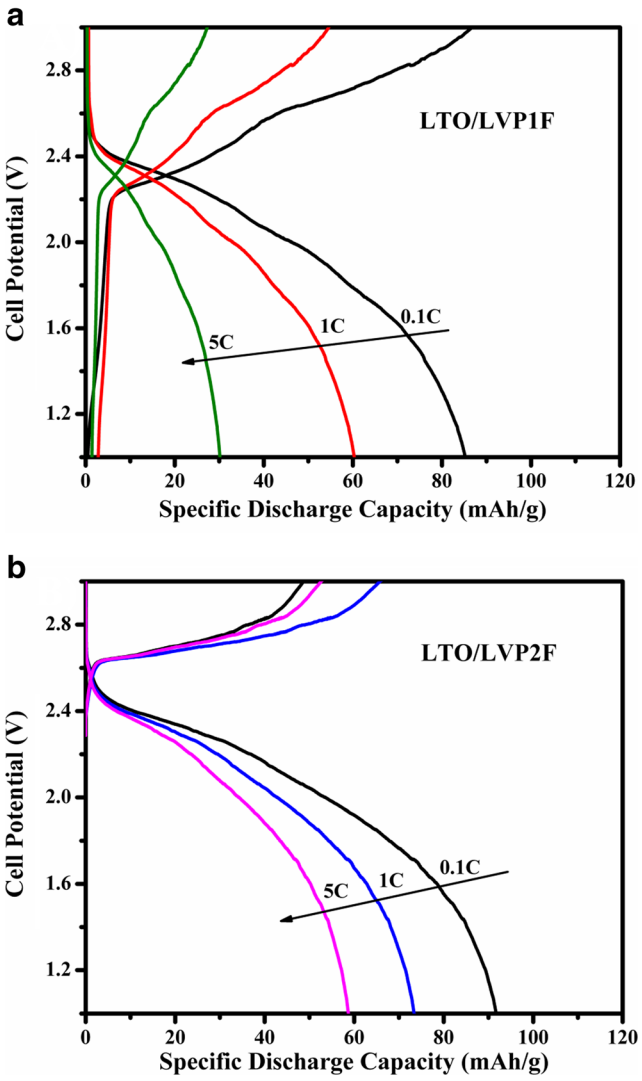


Fig. 6 Rate performances of the LTO/LVP1F (**a**) and LTO/LVP2F (**b**) cells at 0.1 C, 1 C, and 5 C rate

predicted by CV curves. The specific discharge capacities were found to 85 and 92 $mAhg^{-1}$ for LTO/LVP1F and LTO/LVP2F full cells, respectively. These results demonstrate that the mass/capacity balancing factor, i.e., γ plays a significant role in the determination of optimized capacities as it can be seen that full cell configuration with $\gamma = 1.64$ shows better charging/discharging profile as compared to $\gamma = 0.92$.

Figure 6a, b shows the charge/discharge curves of the LTO/LVP1F and LTO/LVP2F cells at different C-rates and the results are summarized in Table 2. It can be observed that LTO/LVP2F delivers the highest discharge capacity at all C-rates as compared to LTO/LVP1F cell. LTO/LVP2F displays an initial discharge capacity of around 92 $mAhg^{-1}$ at 0.1 C rate. Even at high charge/discharge rate of 5 C, LTO/LVP2F exhibits nearly 61% of the discharge capacity delivered at 0.1 C. The decrease in discharge capacity with an increase in current rates can be attributed to increase in polarization potential which consequently hindered lithium-ion kinetics during redox reactions. C. Liu et al. also demonstrate the electrochemical performance of LTO/LVP full battery and reported a discharge capacity of 93 mAh/g at 1 C which is higher than the present work (74 mAh/g for LTO/LVP2F cell) [33]. However, the mass ratio between cathode and anode was kept at 1:1 by Liu et al. which implies an excessive capacity in the anode and, hence, better electrochemical results. K. Zaghib et al. also studied full cell with C-$Li_4Ti_5O_{12}$ as an anode and C-$LiFePO_4$ as a cathode (C-$LiFePO_4$//C-$Li_4Ti_5O_{12}$) instead of C-Li_3V_2 $(PO_4)_3$ at various temperatures with conventional and low-temperature electrolyte and observed that the C-LFP//C-LTO cell with conventional electrolyte exhibits a reversible discharge capacity 1.01 Ah at 0.1 C rate when cycles are between 1.0 and 2.8 V at 25 °C [5].

Cycling results of the rate performance of LTO/LVP1F and LTO/LVP2F cells at 0.1 C, 1 C, and 5 C rate for five charge/discharge cycles are shown in Fig. 7. It can be observed that discharge capacity value for consecutive cycles remains more or less constant for both cells at different C rates. However, LTO/LVP2F cell delivers the highest discharge capacity at each C rate.

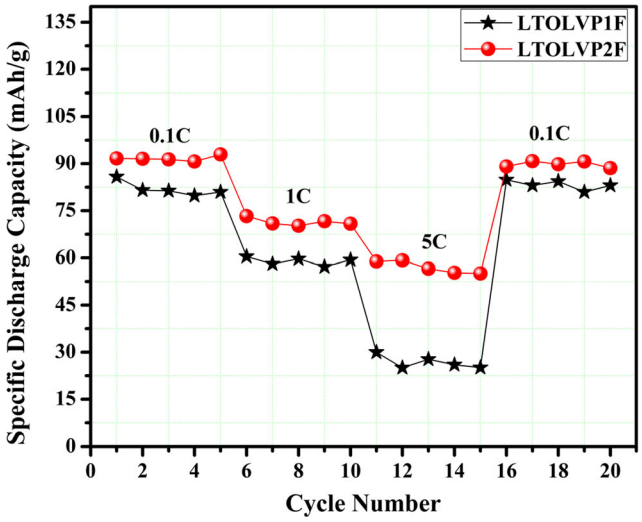


Fig. 7 Cycling results of the rate performance of LTO/LVP1F and LTO/LVP2F cells at 0.1 C, 1 C, and 5 C rate

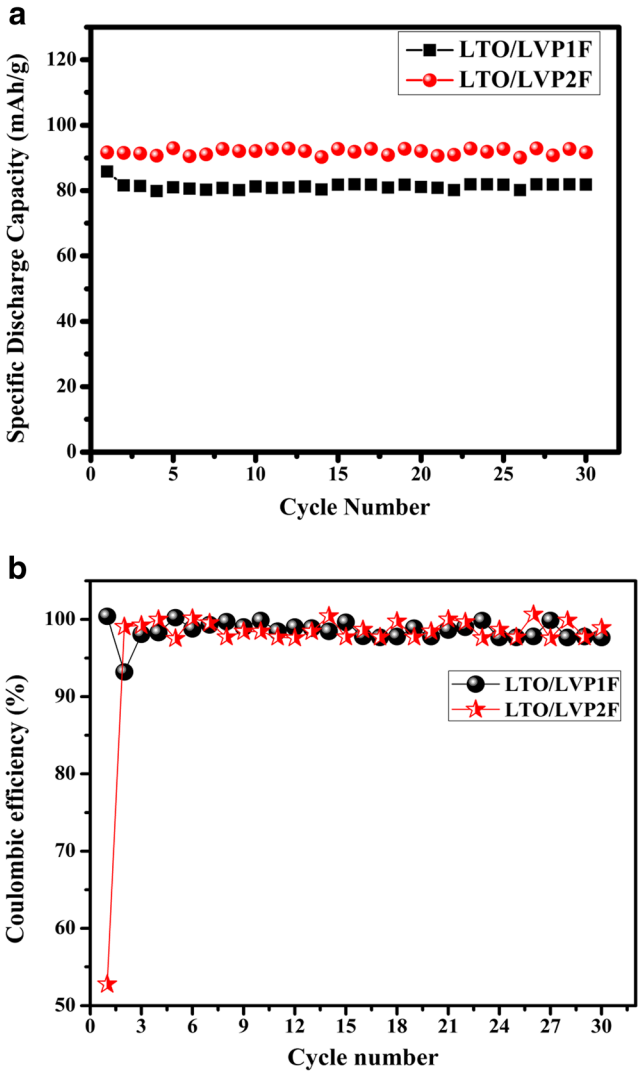


Fig. 8 **a** Cycling performances of LTO/LVP1F and LTO/LVP2F cells at 0.1 C rate for 30 charge/discharge cycles. **b** Coulombic efficiency of full cell for 30 charge/discharge cycles

Table 2 Comparison of electrochemical properties of LTO/LVP1F and LTO/LVP2F cells

Samples	R^a (%)				0.1 C ($mAhg^{-1}$)		
	0.1 C ($mAhg^{-1}$)	0.1 C	1 C	5 C	1st	30th	R^b (%)
LTO/LVP1F	85	100.0	70.5	35.2	85	82	96.4
LTO/LVP2F	92	100.0	80.4	60.8	92	92	100.0

[a] Capacity retention ratios of samples compared to 0.1 C at different C rates

[b] Capacity retention rates of samples compared to the first cycle at 0.1 C

Figure 8a shows the cycling performances of the assembled full cells at 0.1 C rate for 30 charge/discharge cycles, and the cycling capacity retention ratios with respect to the first cycle at 0.1 C rate are summarized in Table 2. The LTO/LVP2F reveals the best cycling performance as it displays an initial discharge capacity 92 $mAhg^{-1}$ at 0.1 C rate and has 100% capacity retention even after 30 charge/discharge cycles at same C rate. Therefore, it can be observed that LTO/LVP2F shows the highest discharge capacity and better rate and cycling performances as compared to LTO/LVP1F cell. Figure 8b displays the Coulombic efficiency of the LTO/LVP1F and LTO/LVP2F cells for 30 charge/discharge cycles at 0.1 C rate. It can be observed that both cells exhibit high Coulombic efficiency except for the first cycle of LTO/LVP2F. The Coulombic efficiency for the subsequent cycles was observed to be more than 95%. These results clearly indicate that the selecting mass/capacity balancing factor, i.e., γ proves to be an important factor to increase the Coulombic efficiency of full cell.

Conclusions

In this work, the effect of mass/capacity balancing factor, i.e., γ has been taken into account in order to study the electrochemical performances of the half-cell as well as full cell configurations of lithium titanate-based anode and lithium vanadium phosphate-based cathodes. XRD results show the formation of pure phases of anode and cathode samples. Electrochemical tests illustrate that the full cell with mass/capacity balancing factor $\gamma = 1.64$, i.e., LTO/LVP2F shows best initial discharge capacity, rate performances at different C rates, and cycling performances at 0.1 C rate. It possessed an initial discharge capacity of around 92 $mAhg^{-1}$ at 0.1 C and retained almost 100% capacity even after 30 charge/discharge cycles at the same rate. Electrochemical tests show that mass/capacity balancing factor γ plays a significant role in optimizing the capacity and rate performances.

Acknowledgements Authors are thankful to the SERB and FIST, Department of Science and Technology (DST), New Delhi, for financial support to carry out this work under the project grant nos. SR/FTP/PS-039/2011 and SR/FST/PSI-164/2011(C), respectively.

References

1. Chen L, Yan B, Xu J, Wang C, Chao Y, Jiang X, Yang G (2015) Bicontinuous structure of $Li_3V_2(PO_4)_3$ clustered via carbon nanofiber as high-performance cathode material of Li-ion batteries. ACS Appl Mater Interfaces 7:13934–13943
2. Shojan J, Chitturi VR, Torres L, Singh G, Katiyar RS (2013) Lithium-ion battery performance of layered $0.3Li_2MnO_3$–$0.7LiNi_{0.5}Mn_{0.5}O_2$ composite cathode prepared by co-precipitation and sol–gel methods. Mater Lett 104:57–60
3. Bhaskar A, Mikhailova D, Kiziltas-Yavuz N, Nikolowski K, Oswald S, Bramnik N, Ehrenberg H (2014) 3d-Transition metal doped spinels as high-voltage cathode materials for rechargeable lithium-ion batteries. Prog solid state ch 42:128–148
4. Hautier G, Jain A, Ong SP, Kang B, Moore C, Doe R, Ceder G (2011) Phosphates as lithium-ion battery cathodes: an evaluation based on high-throughput ab initio calculations. Chem Mater 23: 3495–3508
5. Zaghib K, Dontigny M, Perret P, Guerfi A, Ramanathan M, Prakash J, Mauger A, Julien CM (2014) Electrochemical and thermal characterization lithium titanate spinel anode in C-$LiFePO_4$//C-$Li_4Ti_5O_{12}$ cells at sub-zero temperatures. J. Power Sources 248: 1050–1057
6. Yan B, Chen L, Wang T, Xu J, Wang H, Yang G (2015) Preparation and characterization of $Li_3V_2(PO_4)_3$ grown on carbon nanofiber as cathode material for lithium-ion batteries. Electrochim Acta 176: 1358–1363
7. Yang M, Ren M, Zhu W, Liu W, Zhu C (2015) $Li_3V_2(PO_4)_3$/graphene nanocomposites with superior cycling performance as cathode materials for lithium ion batteries. Electrochim Acta 182: 1046–1052
8. Cui K, Hu S, Li Y (2016) Nitrogen-doped graphene nanosheets decorated $Li_3V_2(PO_4)_3$/C nanocrystals as high-rate and ultralong cycle-life cathode for lithium-ion batteries. Electrochim Acta 210:45–52
9. Rai AK, Thi TV, Gim J, Kim S, Kim J (2015) $Li_3V_2(PO_4)_3$/graphene nanocomposite as a high performance cathode material for lithium ion battery. Ceram Int 41:389–396
10. Luo Y, Xu X, Zhang Y, Pi Y, Zhao Y, Tian X, An Q, Wei Q, Mai L (2014) Hierarchical carbon decorated $Li_3V_2(PO_4)_3$ as a bicontinuous cathode with high-rate capability and broad temperature adaptability. Adv Energy Mater 4:1400107–1400114
11. Gaubicher J, Wurm C, Goward G, Masquelier C, Nazar L (2000) Rhombohedral form of $Li_3V_2(PO_4)_3$ as a cathode in Li-ion batteries. Chem Mater 12:3240–3242
12. Patoux S, Wurm C, Morcrette M, Rousse G, Masquelier C (2003) A comparative structural and electrochemical study of monoclinic $Li_3Fe_2(PO_4)_3$ and $Li_3V_2(PO_4)_3$. J Power Sources 119:278–284
13. Rui X, Yan Q, Skyllas-Kazacos M, Lim TM (2014) $Li_3V_2(PO_4)_3$ cathode materials for lithium-ion batteries: a review. J Power Sources 258:19–38
14. Wei Q, An Q, Chen D, Mai L, Chen S, Zhao Y, Hercule KM, Xu L, Minhas-Khan A, Zhang Q (2014) One-pot synthesized bicontinuous hierarchical $Li_3V_2(PO_4)_3$/C mesoporous nanowires for high-rate and ultralong-life lithium-ion batteries. Nano Lett 14: 1042–1048
15. Huang H, Yin SC, Kerr T, Taylor N, Nazar LF (2002) Nanostructured composites: a high capacity, fast rate $Li_3V_2(PO_4)_3$/carbon cathode for rechargeable lithium batteries. Adv Mater 14:1525–1528
16. Zaghib K, Simoneau M, Armand M, Gauthier M (1999) Electrochemical study of $Li_4Ti_5O_{12}$ as negative electrode for Li-ion polymer rechargeable batteries. J Power Sources 81-82:300–305
17. Guerfi A, Charest P, Kinoshita K, Perrier M, Zaghib K (2004) Nano electronically conductive titanium-spinel as lithium ion storage negative electrode. J Power Sources 126:163–168
18. Guerfi A, Sevigny S, Lagace M, Hovington P, Kinoshita K, Zaghib K (2003) Nanoparticle $Li_4Ti_5O_{12}$ spinel as electrode for electrochemical generators. J Power Sources 119-121:88–94
19. Zhang Z, Cao L, Huang J, Wang D, Meng Y, Cai Y (2013) Temperature effect on spinel Li4Ti5O12 as anode materials for lithium ion batteries. Electrochim Acta 88:443–446
20. Zhang Y, Zhang Y, Huang L, Zhou Z, Wang J, Liu H, Wu H (2016) Hierarchical carambola-like $Li_4Ti_5O_{12}$-TiO_2 composites as advanced anode materials for lithium-ion batteries. Electrochim Acta 195:124–133

21. Wang C, Wang S, Tang L, He YB, Gan L, Li J, Du H, Li B, Lin Z, Kang F (2016) A robust strategy for crafting monodisperse $Li_4Ti_5O_{12}$ nanospheres as superior rate anode for lithium ion batteries. Nano Energy 21:133–144
22. Ni H, Song W, Fan L (2016) Double carbon decorated lithium titanate as anode material with high rate performance for lithium-ion batteries. Prog Nat Sci: Materials International 26:283–288
23. Wang YQ, Gu L, Guo YG, Li H, He XQ, Tsukimoto S, Ikuhara Y, Wan LJ (2012) Rutile-TiO_2 nanocoating for a high-rate $Li_4Ti_5O_{12}$ anode of a lithium-ion battery. J Am Chem Soc 134:7874–7879
24. Lu Y, Zhang F, Zhai B, Zhai S, Xiao Z, An Q, Yu C, Gao S (2016) $Li_4Ti_5O_{12}$ hollow microspheres assembled by nanoparticles as an anode material for lithium-ion batteries. Mater Lett 181:108–112
25. Lin Z, Zhu W, Wang Z, Yang Y, Lin Y, Huang Z (2016) Synthesis of carbon-coated $Li_4Ti_5O_{12}$ nanosheets as anode materials for high-performance lithium-ion batteries. J Alloys Compd 687:232–239
26. Mu D, Chen Y, Wu B, Huang R, Jiang Y, Li L, Wu F (2016) Nano-sized $Li_4Ti_5O_{12}$/C anode material with ultrafast charge/discharge capability for lithium ion batteries. J Alloys Compd 671:157–163
27. Liu T, Ni H, Song WL, Fan LZ (2015) Enhanced electrochemical performance of $Li_4Ti_5O_{12}$ as anode material for lithium-ion batteries with different carbons as support. J Alloys Compd 646:189–194
28. Zhao B, Ran R, Liu M, Shao Z (2015) A comprehensive review of $Li_4Ti_5O_{12}$-based electrodes for lithium-ion batteries: the latest advancements and future perspectives. Mater Sci Eng R-Rep 98:1–71
29. Ohzuku T, Ueda A, Yamamoto N (1995) Zero-strain insertion material of Li $[Li_{1/3}Ti_{5/3}]O_4$ for rechargeable lithium cells. J Electrochem Soc 142:1431–1435
30. Arora P, White RE, Doyle M (1998) Capacity fade mechanisms and side reactions in lithium-ion batteries. J Electrochem Soc 145: 3647–3667
31. Yan H, Zhang D, Guo G, Wang Z, Liu Y, Wang X (2016) Hydrothermal synthesis of spherical $Li_4Ti_5O_{12}$ material for a novel durable $Li_4Ti_5O_{12}$/$LiMn_2O_4$ full lithium ion battery. Ceram Int 42: 14855–14861
32. Zhao B, Ran R, Liu M, Shao Z (2015) A comprhensive review of $Li_4Ti_5O_{12}$-based electrodes for lithium-ion batteries: the latest advancement and future prospectives. Mater Sci Eng R 98:1–71
33. Liu C, Wang S, Zhang C, Fu H, Nan X, Yang Y, Cao G (2016) High power high safety battery with electrospun Li3V2(PO4)3 cathode and Li4Ti5O12 anode with 95% energy efficiency. Energy Storage Materials 5:93–102

Accepted Manuscript



Title: Estimation of intrinsic work function of multilayer graphene by probing with electrostatic force microscopy

Author: <ce:author id="aut0005" author-id="S0169433217300478-cff32820963c8db2b6505f5cc7403dbd"> Anshika Singh<ce:author id="aut0010" author-id="S0169433217300478-944d7a105f2b801f9d8f16e8975ddfb6"> Puspendu Guha<ce:author id="aut0015" author-id="S0169433217300478-30ff243546280ac802b343ee2767aacf"> Amrish K. Panwar<ce:author id="aut0020" author-id="S0169433217300478-960b97d76dbe835b49f1b0486eae90eb"> Pawan K. Tyagi

PII: S0169-4332(17)30047-8
DOI: http://dx.doi.org/doi:10.1016/j.apsusc.2017.01.047
Reference: APSUSC 34842

To appear in: *APSUSC*

Received date: 19-10-2016
Revised date: 23-12-2016
Accepted date: 6-1-2017

Please cite this article as: Anshika Singh, Puspendu Guha, Amrish K.Panwar, Pawan K.Tyagi, Estimation of intrinsic work function of multilayer graphene by probing with electrostatic force microscopy, Applied Surface Science http://dx.doi.org/10.1016/j.apsusc.2017.01.047

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

Estimation of intrinsic work function of multilayer graphene by probing with electrostatic force microscopy

Anshika Singh[1],Puspendu Guha[2,3], Amrish K. Panwar[1], Pawan K. Tyagi[1#]

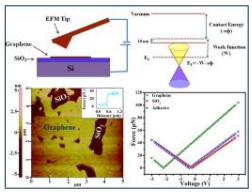
[1]Department of Applied Physics, Delhi Technological University, Delhi, 110042

[2]Institute of Physics, SachivalayaMarg, Bhubaneswar-751005, India

[3]Homi Bhabha National Institute, Training School Complex, Anushakti Nagar,Mumbai-400085, India

[#]pawan.phy@dce.edu, pawankumartygai@gmail.com

Graphical abstarct



ACCEPTED MANUSCRIPT

Highlights

- The surface contact potential and work function of FLG films on SiO_2 and effect of adhesive have been studied by employing an EFM technique.
- EFM study on charged graphene surface has been performed by using dc tip-bias.
- The infinitesimal variation of the surface contact potential on graphene sheet has been probed through the histogram analysis of force images.
- Surface potentials at the edges of graphene sheet found to have sharp variation, which might be attributed to the unsaturated bonds and vacancies at these sites.
- By adopting this method intrinsic work function of graphene is determined to be 4.52 eV.

Abstract

In present study, electrostatic force microscopy (EFM) is used to estimate the intrinsic work function of few layer graphene (FLG) transferred on SiO_2 (300nm)/Si (500 μm) substrate. This FLG has been prepared by using the mechanical exfoliation technique. In exfoliated FLG,adhesive residues are always left from scotch tape on its surface. These residues as well as SiO_2 substrate could modify the work function due to the formation of dipoles on the surface. Taking the effect of adhesive into account, FLG ispre-charged and then scanned with a tipbiased with dc voltage.Intrinsic work function of FLG is determined and found to be 4.52 ± 0.1eV.

Keywords: Electrostatic Force Microscopy; Surface potential; Few layer graphene; Mechanical exfoliation.

ACCEPTED MANUSCRIPT

1.Introduction:

Recently, graphene has attracted huge research interest due to its unique 2-dimensionalstructure exhibiting ballistic transport at submicron length scales [1]and extraordinary electronic[2], mechanical[3], optical[4], chemical[5]and thermal [6] properties. Since, graphene has all the atoms on itssurface, interaction with environment plays a vital role[7] andcan easily influence the intrinsic mobility as well as other electronic properties[8-9]such assurface potential, charge distribution and work function. Recently, by employing standard electrostatic force microscope (EFM) technique local electrostatic force on the graphene surface has been probed [8,10,11] . Work function of single layer graphene was reported to be ~ 4.55 ± 0.02eV[12]which is equivalent to that of graphite (~4.6 eV)[13] and found to be increasing as number of layers increase[14].However, Panchal.*et.al* reported that work function of bi-layer graphene (~ 4.44 ± 0.02 eV) is lower than that of single layer[12] which was grown through sublimation of Si-terminated face of an on-axis *4H*-SiC (0001) at 2000 °C. Graphene grown on SiC substrate was reported to possess different electronic properties compared to CVD-grown or mechanically exfoliated.Recently, the surface contact potential difference (CPD)between tip and FLG was also reported to increase as layer number increased[15] and found to be greatly influenced by the adsorbent on the surface of FLG[8] as well as FLG/substrate interface[16].Therefore, it is necessary to measure the intrinsic surface potential of FLG accurately even if adhesive and adsorbent are present.

In presentstudy, we aim to estimate the intrinsic CPD and work function of FLG prepared by scotch tape exfoliation method. In order to nullify the influence of residual adhesive, we have first pre-charged FLG surface and then scanned with biased tip. Variation in CPD of FLG, adhesive as well as SiO_2 substrate has been studied.We found that adopted methodology

ACCEPTED MANUSCRIPT

provides an excellent way to determine intrinsic work function of FLG and demonstrates thatwork functioncan be determinedeven if adsorbentsare present on the surface of FLG.

2. Experimental details:

FLGs were prepared by mechanical exfoliation method.Briefly, flakes of FLG were detached from HOPG (Highly oriented pyrolytic graphite)using adhesive tape(Scotch tape 3M MAGIC TAPE 811supplied by SPI).These flakeswerecleaved repeatedly until thin layer of graphene was left on the tape and thentransferred onto SiO_2 (300nm)/Si (500 μm) substrate. Obtained FLGs were optically identified and probed with EFM. In order to carry out EFM measurements, scanning probe microscope (Park, Model XE-100) with gold coated Si tip (NCSTAu) was used where resonance frequency and set point of the conductive tipwerekept at 137 kHz and 21.04 nm,respectively.In order to pre-charge the FLG, charge injection was performed by following the procedure as reported earlier[9]. In the process,image of FLG was taken with the grounded tip in contact modeand then tip was brought into contact with FLG as shown in Fig. 1(b). Thereafter, a bias V_{inj} (-1.5 V)was applied to the tip for 4 minutes followed by retraction of tip to 50 nm. FLG gotchargedin such a way that charge resides on its surface whereas in case of insulating SiO_2 or adhesive, charge resides in the interior. A dc bias was applied to the tip and then force, amplitude and phase images were recordedat different tip-bias-voltage (V_{EFM}) in range 3 V to -3 V as shown in Fig. 1(c).By controlling thetip bias,electrostatic force caused due to differencein the Fermi energyof tip and FLG, was nullified and local surface potential contact difference (LCPD) was mapped out.This procedure was applied to FLG, pristine as well as SiO_2/Si substrate affected by adhesive residues left from scotch tape.

3. Results and discussion:

ACCEPTED MANUSCRIPT

Typical EFM topography and force image of FLGtransferred onSiO_2 (300nm)/Si (500 μm) substrate are shown in Fig. 2(a) and Fig. 2(c), respectively. The line profileon topography image is drawn and plotted in Fig. 2bto estimate the thickness of graphene layer. The height difference between SiO_2 substrate and graphene layer represents the thickness of FLG and determined to be ~1.75 nm.Since, during force imaging tip was lifted up to 50 nm, it is presumed that electrostatic force dominated while vander Waals force was minimized. The force imaging was performed keeping tip bias at -0.5 Vas show in Fig. 2(c). In the figure,equipotential areascan be distinguishedas darker regioncorresponds to SiO_2/Si substrate and brighter one to FLG.At the edge, bright contrast reveals that the surface potential is low. This might have happened dueto altered C=C bonds either by folded wrinkle or by the attachment of functional groups at this site[17]. It furtherreveals that at this site, work functionwas locallyincreased.An increase in electrostatic force as well as corresponding reduction insurface potential at theedges and corner sites might be due to defects, adsorbents or large amount of unsaturated bonds. It is well known fact that the local work function of graphene layers increases as concentration of unsaturated bonds or vacancy increases [18]. Furthermore, charge exchange at graphene/ SiO_2 interface has been reported to influence theCPDresulting from the screening of interlayer coupling[17,19]. Previously, it was reported that Dirac point, where electron and hole conduction meet, significantly influenced with surface characteristic and dielectric constant of the substrate. It was reported that surface characteristic (untreated or functionalized) of SiO_2 lead to effective p-or n-type doping in graphene[20]. Since, in present case FLG was first pre-charged and then scanned with biased tip. Thin layered SiO_2 might increase the probability of leakage of the charge from graphene to SiO_2/Si substrates. As consequence, CPD as well as work function might vary.Therefore, work function of FLG might depend on the thickness of SiO_2. However, if FLG

ACCEPTED MANUSCRIPT

was not pre-charged then it will be unable to influence the work function.On further investigation,force versus distance curve along the line profile as markedin force image has been plotted in Fig. 2(d). The force along the line was found to be32 pN on FLG whereas 0.79 pN on SiO_2 substrate.

The CPD of FLGwas also mapped by keeping different V_{tip} as mentioned at the bottom of the panels of Figs. 3(a-h). In force images, bright areas are clearly visible which corresponds to graphene whereas darker area corresponds to SiO_2 substrate. Force image reveals that as tip-bias voltage (V_{EFM})decreases,both brightnessas well as darkness of corresponding regions is reduced. When $V_{tip}$was taken in the range -1.5 V to -2 V, force image could not be obtained. It was expected because FLG was pre-charged with -1.5 V and takingthis into account, we found that variation in the surface potential is in range of 0 to -0.5 V. Contrarily, at V_{tip}= -2.5 V both the regions became brighter. As evident from Figs. 4&5phase and amplitude were found to be more sensitive than the force.Unlike force image, both phase as well as amplitude images were obtained at V_{tip} applied in range of -1 V to -2 V and variation in CPD was more accurately probed. In phase and amplitude images,edge sites were found to exhibit large variation in CPDas evident by brighter edge sites of FLG. It can be concluded that in order to examine the variation in CPD, minimum required potentialis 0.5 V. Magnitude of electrostatic force gets minimized when V_{tip} is equivalent to CPD as shown in Fig.6 (b). Large shift in CPD of bare SiO_2 and FLG transferred on SiO_2 was observed whereas it is minimal in case of bareand contaminated (adhesive) SiO_2. As evident in Fig.6, force minima at -2.2V, -0.4V and -0.2V were observed for FLG (5 layers, marked by line 1), contaminated and bare SiO_2, respectively. Furthermore,previously it was reported that electric field may exist at the surface due to adhesive residuesor ambient environment which can influence CPD of FLG. Performed study further

ACCEPTED MANUSCRIPT

confirms that adsorbents affect both CPD as well as work function, as evidenced byincrease in negative values of force minima. Previously, it was reported that both adsorbates and substrate could modify the surface potential of graphene by an amount (V_a). Thus, the resultant surface contact potential difference (V) can be expressed as a two components, $\Delta V = \Delta V_{CPD} + \Delta V_a$[21, 22],where V_{CPD} is the intrinsic CPD of the graphene. In present case, after eliminating the effect of substrate and adhesive, force minima for intrinsicFLG was determined to be -1.8 V.In present case, FLG was already biased with -1.5 Vtherefore intrinsic CPD between tip and FLG as well as adhesive and bare substrate would be - 0.3eV, 1.1 eV and 1.3 eV, respectively. After estimating intrinsic CPD, work function of FLG was calculated by using the relation[17,21,23]:$\Delta V_{CPD} = \frac{\phi_{tip} - \phi_{graphene}}{-e}$where $\Phi_{graphene}$ and Φ_{tip} are the work function of graphene and tip, respectively. Taking the previously reported value of work function for gold coated tip4.82± 0.08 eV[23], intrinsic work function of FLG was determined to be 4.52 ± 0.1 eV. Similarly, this valueis found 6.12 ± 0.1 eV for SiO_2. We have also observed that work function of FLG was increased with number of layers.As marked in Fig 6(a), three different region in topographical image having thickness 1.75 nm (5 layers), 2.5 nm (7 layers) and 4 nm (12 layers) exhibit corresponding force minima exist at -2.20 V, -2.15 V and -2.10 V, respectively as shown in Fig. 6(b). Work functions of these layers are determined to be 4.52 ± 0.1 eV, 4.57 ± 0.1 eV and 4.62 ± 0.1 eV. These observations evidence that work function of FLG is layer dependent.

Underlying mechanism of the performed experiment can be understood as: initially graphene was pre-charged and Fermi level of graphene was lifted-up by an amount eV_{inj}. When the tip was brought closer to the graphene surface, an electrical force was generated between tip and sample. This happened due to mismatch in their Fermi energy levels[24]. The initial mismatch

ACCEPTED MANUSCRIPT

ofΦ_{tip} and Φ_G,position of vacuum and Fermi levelsis depicted inFig.7 (a) when they are not electrically connected. As shown in Fig. 7 b, the pre-charging of graphene lifted up its Fermi level.. During the scanning, tip and pre-charged graphene surface was close enough for electron tunneling to occur. Consequently, their Fermi levels got aligned whereas vacuum level of graphene went down by an amount V_{CPD}. An electrical force generated in the contact area, due to CPD was nullified by applying external bias (V_{DC}) to the tip. The amount of total applied bias between tip and graphene (V_{inj}- V_{tip}) which nullifies the electrical force is equal to the difference in the work function of tip and graphene.

Lateral variations in the intrinsic work function of FLG might have occurred due to in plane screening of charge with impurities present on its surface.Therefore, these variations were further examined through the close views of both topography and force images as well as histogram of force images for FLG transferred on SiO_2, contaminated and bare SiO_2, as shown in Fig.8. No significant variation has been observed in their topographical images. This suggests that contaminations if present, lying on the SiO_2 substrate or on graphene surfaces are very small to be imaged at this resolution. However,force images and corresponding histogramsof FLGand bare SiO_2 showsignificant variations. The histogram of both FLG and contaminated SiO_2 provide better comparison as shown in Figs. 8 (g & h). In case of FLG, we observes an asymmetric distribution of the surface potential as shown in Fig. 8 (g), whereasFigs.8 (h&i) exhibit relatively symmetric distribution for contaminated and bare SiO_2, respectively. The variations in the surface potential on FLG as shown in Fig. 8d can be attributed to partial screening of the charge through impurities[24] or to the doping effect of environment or adhesive.

4. Conclusion:

ACCEPTED MANUSCRIPT

We have determined the intrinsic work function of FLG at ambient conditions. Unlike previously reported procedures, we have first pre-charged the FLG and then applied dc tip-bias.The adopted procedure has been demonstrated as an efficient method for the estimation of surface properties such as surface contact potential and intrinsic work function of few layer graphene. In this study, intrinsic work function of FLG films on SiO_2 have been determined by employing an EFM technique where instead of ac tip-bias, dc bias was used. Intrinsic work function of FLG is determined to be 4.52 ± 0.1 eV in well agreement with reported value. Surface potential at the edges of FLG sheet were found to have sharp variation attributed to the unsaturated bonds and vacancies at these sites.

Acknowledgement

We thank the financial support from DST-SERB, fast track Program (SR/FTP/PS-055/2012).

5. References:

1. X. Du, I. Skachko, A. Barker, E.Y. Andrei, Approaching ballistic transport in suspended graphene Nat. Nanotechnol. 3 (2008) 491-494.
2. A.H.C. Neto, F. Guinea, N.M.R. Peres, K.S. Novoselov, A.K. Geim, The electronic properties of graphene, Rev. Mod. Phys. 81 (2009) 109-162.
3. C. Lee, X.D. Wei, J.W. Kysar, J. Hone, Measurement of the elastic properties and intrinsic strength of monolayer graphene, Science 321(2008) 385-388.
4. F. Bonaccorso, Z. Sun, T. Hasan, A.C. Ferrari, Graphene photonics and optoelectronics, Nature Photon. 4 (2010**)** 611-622.
5. J.D. Fowler, M.J. Allen, V.C. Tung, Y. Yang, R.B. Kaner, B.H. Weiller, Practical chemical sensors from chemically derived graphene, ACSNano 3 (2009) 301-306.

ACCEPTED MANUSCRIPT

6. A. A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, C. N. Lau, Superior thermal conductivity of single-layer graphene, Nano Lett. 8 (2008) 902-907.

7. J. Sabio, C. Seoanez, S. Fratini, F. Guinea, A. H. Castro Neto, F. Sols, Electrostatic interactions between graphene layers and their environment, Phys. Rev. B, 77 (2008) 195409.

8. J. Moser, A. Verdaguer, D. Jimenez, A. Barreiro, A. Bachtold, The environment of graphene probed by electrostatic force microscopy,Appl. Phys. Lett. , 92 (2008**)** 123507-1-3.

9. A. Verdaguer, M. Cardellach, J.J. Segura, G.M. Sacha, J. Moser, Charging and discharging of graphene in ambient conditions studied with scanning probe microscopy Appl. Phys. Lett. 94 (2009)233105-1-3.

10. O. Kazakova, V. Panchal, T.L. Burnett, Epitaxial graphene and graphene–based devices studied by electrical scanning probe microscopy, Crystals 3 (2013) 191-233.

11. S. Lilliu, C. Maragliano, M. Hampton, M. Elliott, M. Stefancich, M. Chiesa, M.S. Dahlem, J.E. Macdonald, EFM data mapped into 2D images oftip-sample contact potential difference and capacitance second derivative, Sci. Rep.3 (2013) 1-7.

12. V. Panchal, R. Pearce, R. Yakimova, A. Tzalenchuk, O. Kazakova, Standardization of surface potential measurements of graphene domains,Nature 3 (2013) 1-8.

13. T. Takahashi, H. Tokailin, T. Sagawa, Angle-resolved ultraviolet photoelectron spectroscopyof the unoccupied band structure of graphite, Phys. ReV. B32 (1985) 8317-8324.

ACCEPTED MANUSCRIPT

14. H. Hibino, H. Kageshima, M. Kotsugi, F. Maeda, F.-Z. Guo, Y. Watanabe, Dependence of electronic properties of epitaxial few-layer graphene on the number of layers investigated by photoelectron emission microscopy,Phys. ReV. B79 (2009) 125437-1-7.

15. S.S. Datta, D.R. Strachan, E.J. Mele, A.T.C. Johnson, Surface potentials and layer charge distributions in few-layer graphene films, Nano Lett. 9 (2008) 7-11.

16. Y. Shi, X. Dong, P. Chen, J. Wang, L.J. Li, Effective doping of single-layer graphene from underlying SiO2 substrates, Phys. Rev. B 79 (2009**)** 115402-1-4.

17. J. Li, X. Qi, G. Hao, K. Huang, J. Zhong, Surface potential of graphene oxide investigated by kelvin probe force microscopy, Fullerenes, Nanotubes, Carbon Nanostruct. 23 (2015) 777-781.

18. W. Wang, Z. Li.,Potential barrier of graphene edges,J. Appl. Phys. 109 (2011**)** 114308-1-6.

19. N.J. Lee, J.W. Yoo, Y.J. Choi, C.J. Kang, D.Y. Jeon, D.C. Kim, S. Seo, H.J. Chung, The interlayer screening effect of graphene sheets investigated by Kelvin probe force microscopy,Appl. Phys. Lett. 95 (2009) 222107-1-3.

20. J. Park, W.H. Lee, S. Huh, S.H. Sim, S.B. Kim, K. Cho, B.H. Hong, K.S. Kim, Work function engineering of graphene electrodes by self assembled monolayers for high performance organic field effect transistor, J. Phys. Chem. Lett., 2 (2011) 841-845.

21. W. Melitz, J. Shen, A.C. Kummel, S. Lee, Kelvin probe force microscopy and its application, Surf. Sci. Rep. 66(2011) 1-27.

22. G.L. Hao, X. Qi, J. Li, L.W. Yang, J.J. Yin, F. Lu, J.X. Zhong, Surface potentials of few-layer graphene films in high vacuum and ambient conditions, Solid State Commun.,151 (2011) 818-821

ACCEPTED MANUSCRIPT

23. Y.-J. Yu, Y. Zhao, S. Ryu, L.E. Brus, K.S. Kim, P. Kim,Tuning the graphene work function by electric field effect, NanoLett. 9 (2009) 3430-3434.

24. D. Ziegler, P.Gava, J. Guttinger, F. Molitor, L.Wirtz,M.Lazzeri, A.M.Saitta,A. Stemmer, F. Mauri, C. Stampfer, Variations in the work function of doped single- and few-layer graphene assessed by Kelvin probeforce microscopy and density functional theory,Phys. Rev. B83 (2011) 235434-1-7.

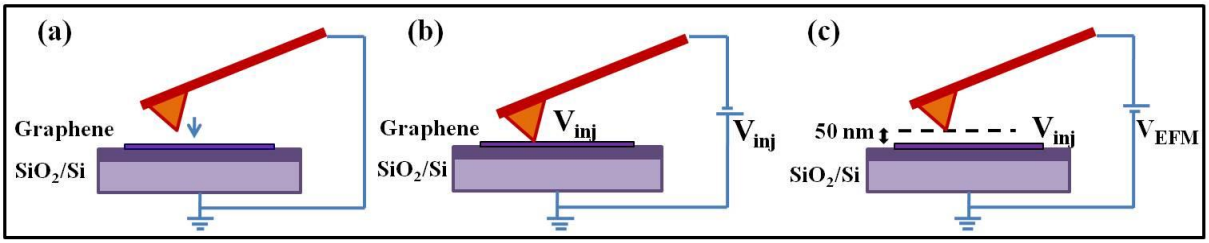


Fig.1. (a-c) Schematic of the charge injection procedure and EFM measurement

ACCEPTED MANUSCRIPT

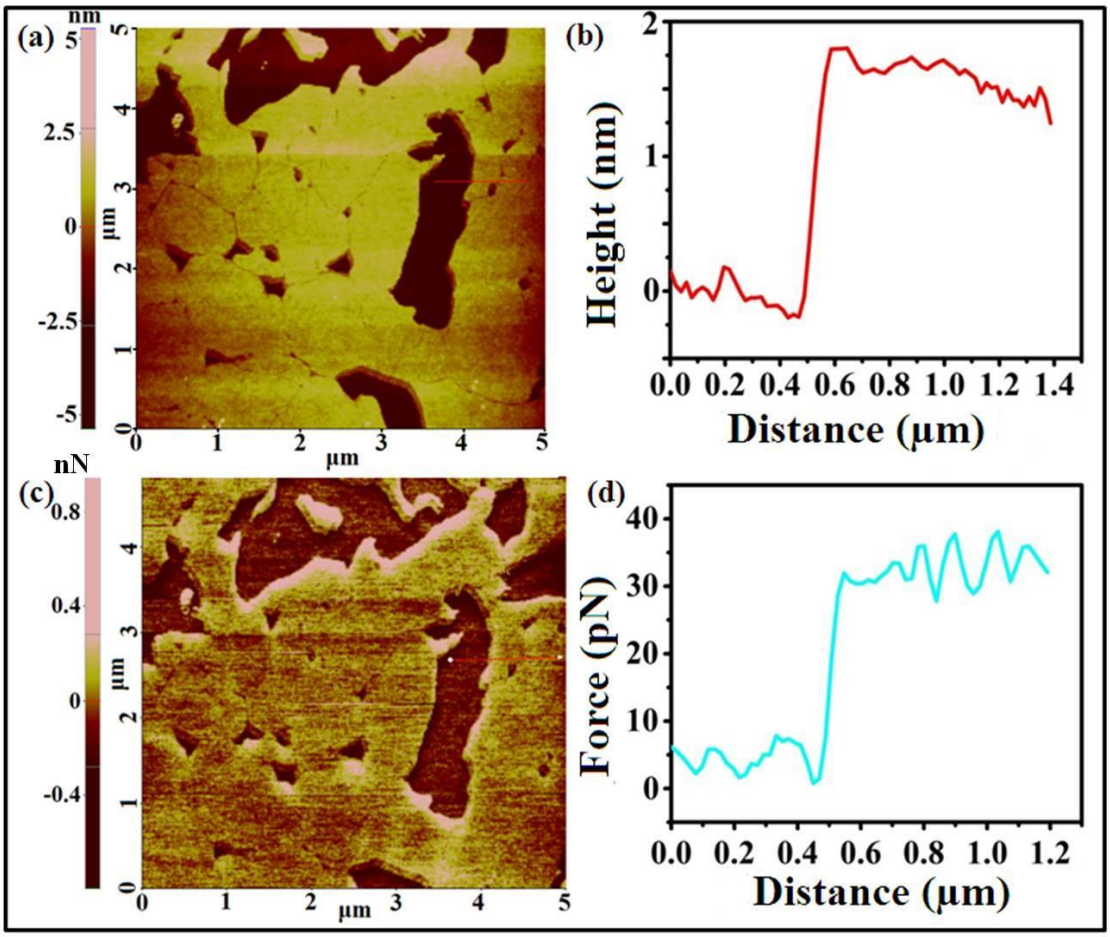


Fig.2. (a) EFM topography image of FLG on SiO_2/Si substrate, (b) Heightprofile along line drawn on the topography image, (c) force image and (d) corresponding force profile.

ACCEPTED MANUSCRIPT

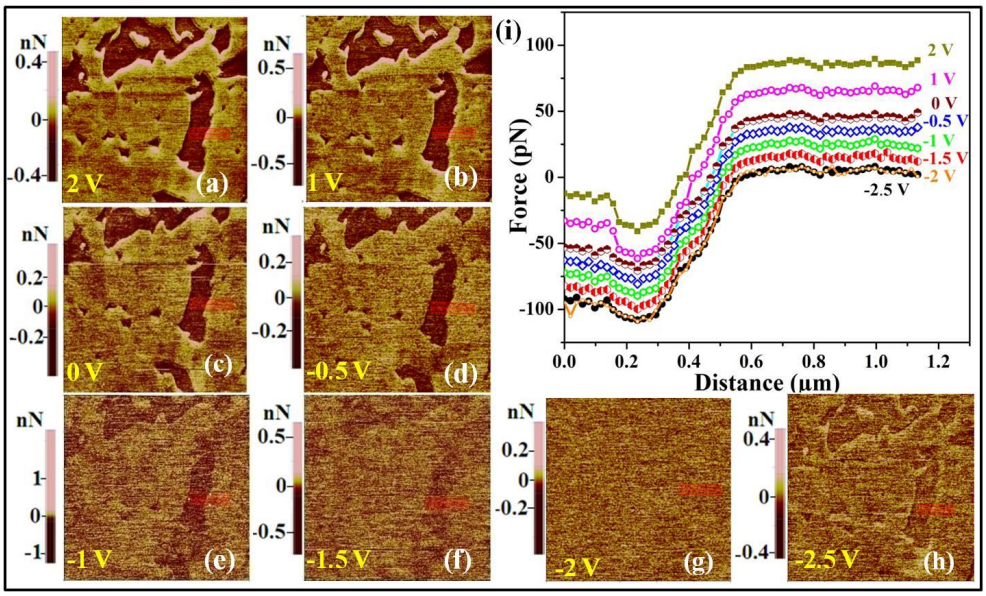


Fig.3. Force image of FLG (5x5 μm area) (a) 2V,(b) 1V, (c) 0V, (d) -0.5V, (e) - 1V, (f) - 1.5V, (g) – 2V, (h) -2.5V tip voltage and (i) correspondingaverage forceprofile.

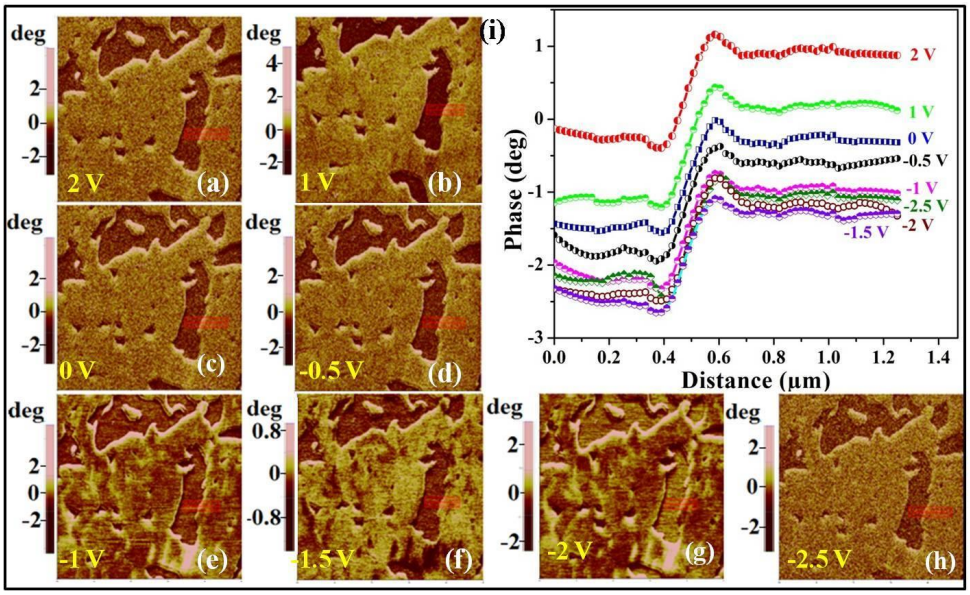


Fig.4. Phase image of FLG (5x5 μm area) (a) 2V,(b) 1V, (c) 0V, (d) -0.5V, (e) -1V, (f) -1.5V, (g) -2V, (h) -2.5V tip voltage and (i) corresponding average phase profile.

ACCEPTED MANUSCRIPT

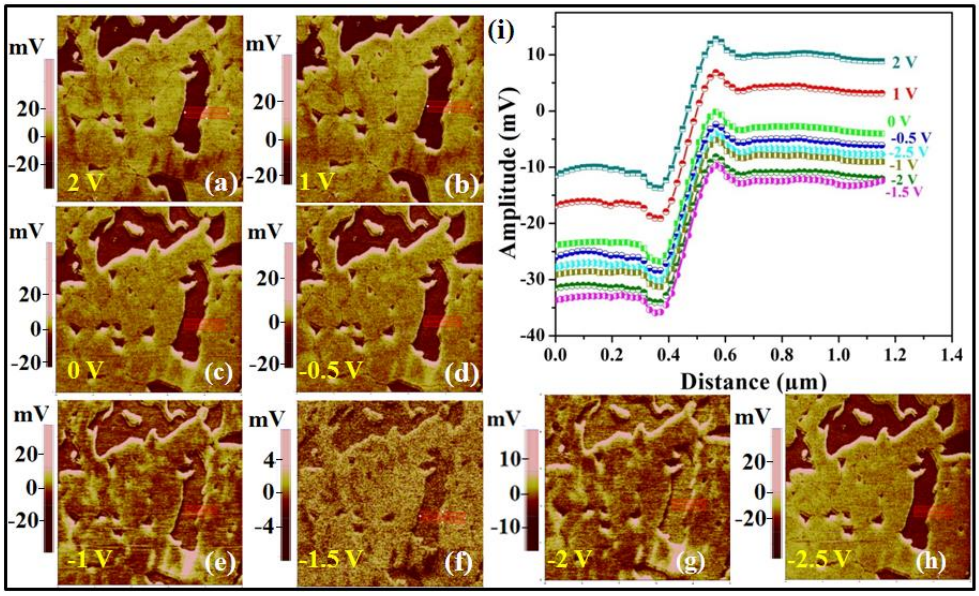
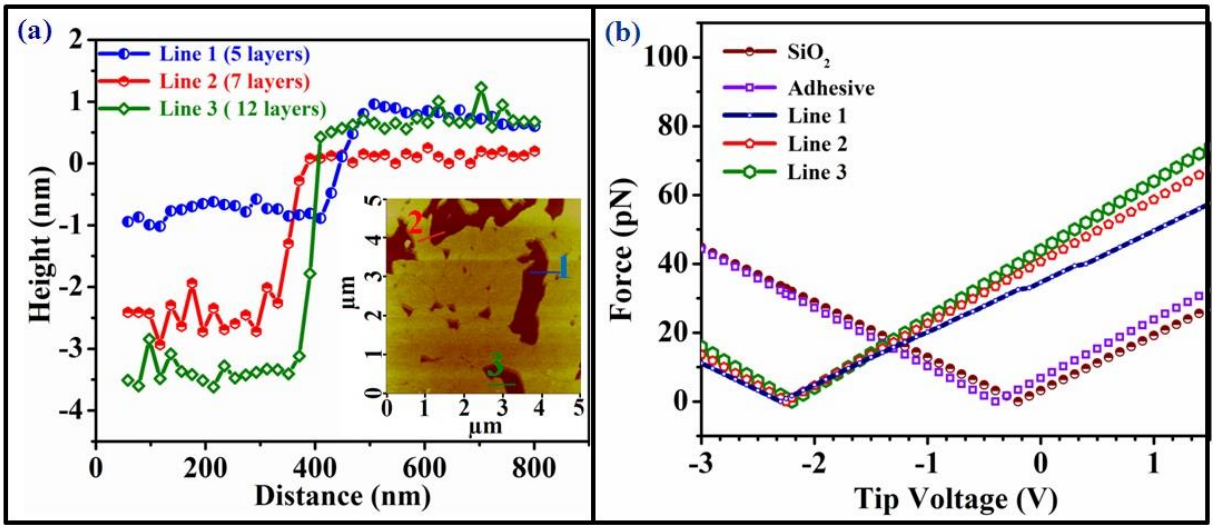


Fig.5. Amplitudeimage of FLG (5x5 µm area) (a) 2 V,(b) 1 V, (c) 0 V, (d) - 0.5 V, (e) - 1 V, (f) - 1.5 V, (g) - 2 V, (h) - 2.5 V tip voltage and (i) corresponding average amplitude profile.



ACCEPTED MANUSCRIPT

Fig.6.(a) Height profile representing the number of layers in as marked lines in topography image shown in lower set of panel, (b) Electrostatic force experienced between tip and FLG at different tip bias on bare, contaminated with tape adhesive SiO_2, FLG transferred on SiO_2 substrate and different marked region.

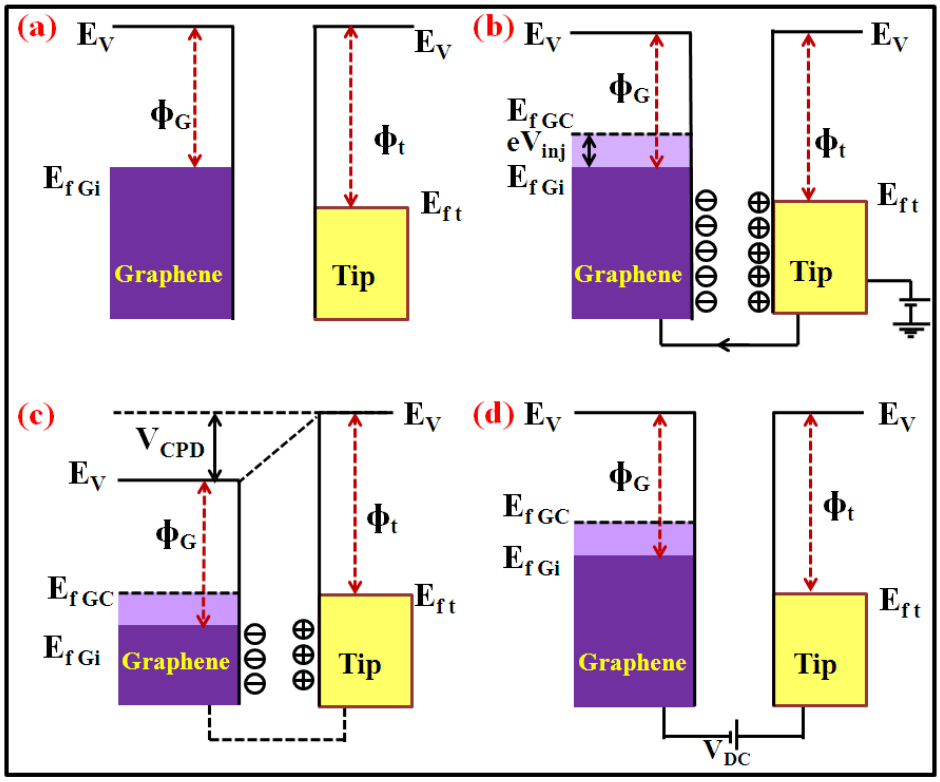


Fig. 7. Electronic energy levels of FLG and tip for different cases: when (a) tip and FLG are electrical insulated, (b) FLG pre-charged through tip, (c) tip at height of 50 nm, and (d) V_{DC} applied to tip.

ACCEPTED MANUSCRIPT

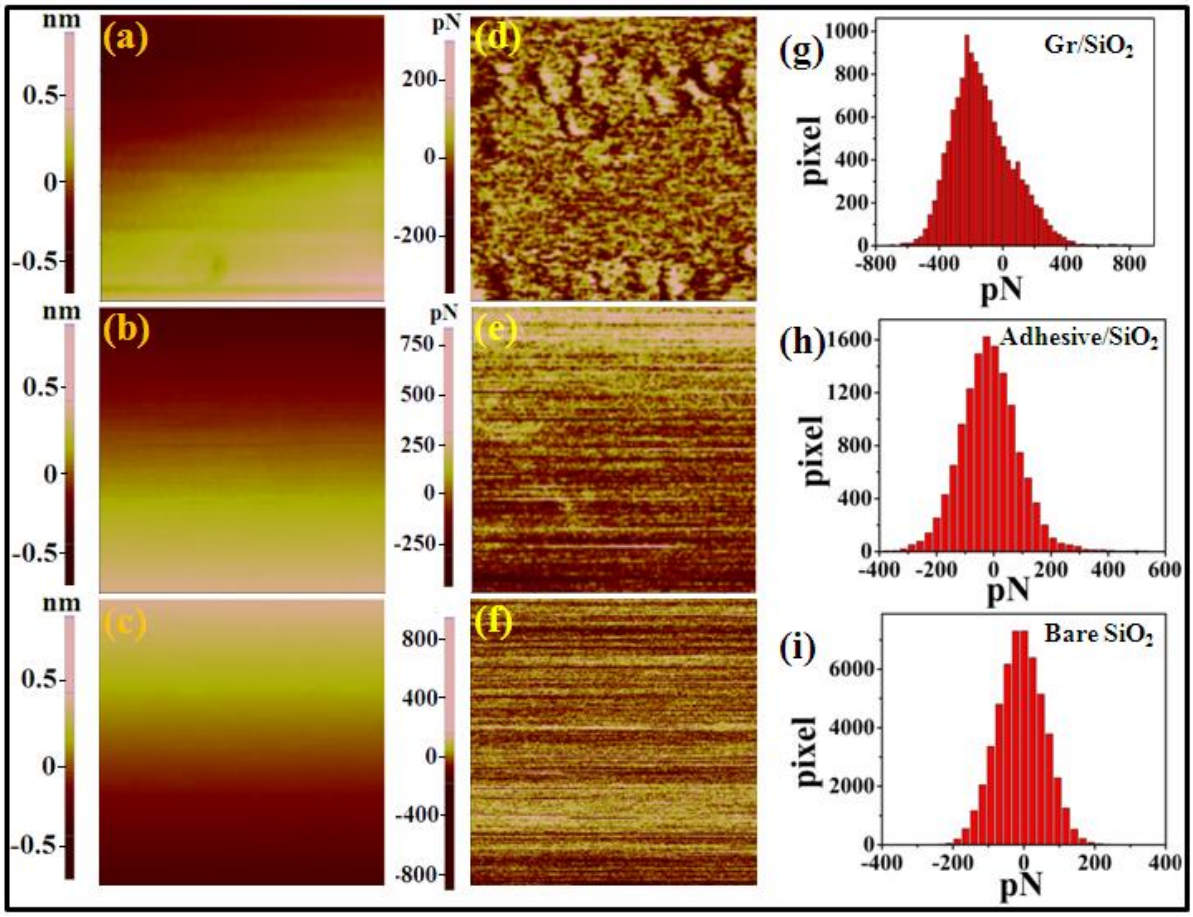
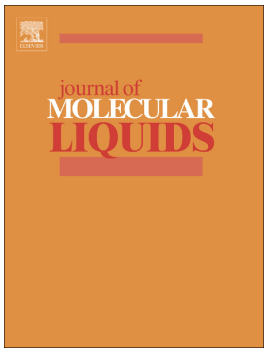


Fig.8. (a-c) Topographical image(900 ×900 nm) taken of FLG,contaminated and bare SiO_2, respectively.Corresponding histograms shown in panels (g-i).

Accepted Manuscript



Investigation of charge-separation/change in dipole moment of 7-azaindole: Quantitative measurement using solvatochromic shifts and computational approaches

Mohan Singh Mehata, Ajay K. Singh, Ravindra Kumar Sinha

PII: S0167-7322(16)32761-1
DOI: doi: 10.1016/j.molliq.2017.01.091
Reference: MOLLIQ 6893

To appear in: *Journal of Molecular Liquids*

Received date: 15 September 2016
Accepted date: 24 January 2017

Please cite this article as: Mohan Singh Mehata, Ajay K. Singh, Ravindra Kumar Sinha , Investigation of charge-separation/change in dipole moment of 7-azaindole: Quantitative measurement using solvatochromic shifts and computational approaches. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Molliq(2017), doi: 10.1016/j.molliq.2017.01.091

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

Investigation of Charge-Separation/Change in Dipole Moment of 7-Azaindole: Quantitative measurement using Solvatochromic shifts and Computational approaches

Mohan Singh Mehata[a,*], Ajay K. Singh[b] and Ravindra Kumar Sinha[a,c]

[a]Laser-Spectroscopy Lab, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi-110042, India.

[b]Radiation & Photochemistry Division, Bhabha Atomic Research Centre, Trombay, Mumbai 400 085, India.

Corresponding author: E-mail: msmehata@gmail.com; mohan.phy@dce.edu

[c]*Present address: CSIR-Central Scientific Instruments Organisation, Sector 30, Chandigarh-160 030.*

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

Solvatochromic and computational approaches were applied to 7-azaindole (7-AI). The vertical transitions together with the bright singlet states and HOMO-LUMO were obtained from the DFT/TDDFT quantum chemical calculations. The ground and excited electronic states dipole moments hence the change in the dipole moment were investigated by measuring the steady-state absorption and fluorescence spectra in various solvents and by using DFT/TDDFT calculations. The dipole moment in the photo-excited state is significantly higher than the ground state, as obtained both experimentally and theoretically. The higher values of excited state dipole moment than the ground state dipole moment demonstrating a significant charge separation in the photo-excited state of 7-AI, which in turn depends on the solvent polarity and hydrogen bonding ability. Also, the nearly linear relation observed between the fluorescence maximum/Stokes shift and solvent polarity indicates sensing ability of 7-AI towards organic solvents**.**

Keywords: 7-azaindole, absorption and fluorescence, Stokes shift, dipole moment, charge separation, solvents, DFT/TDDFT calculations.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1. Introduction

7-Azaindole (7-AI) has been considered as a model molecule for the study of DNA because there are some similarities between the 7-AI dimer and adenine-thymine (A-T) as well as guanine-cytosine (G-C) base pairs of DNA [1-4], as hydrogen bonded bridge links all these complexes. The two N ⋯ H – N hydrogen bonds in the 7-AI dimer being formally identical with one of those in both the A-T and G-C pairs. Thus, the photo-induced tautomeric behavior of 7-AI and DNA base pairs could serve as a model for the possible mutagenic behavior in genetics [3]. The functional groups of 7-AI are in such a positions that photo-induced tautomerization results from a concerted double proton transfer from one group to the other in hydrogen-bonded 7-azaindole dimer [1,2]. The studies of double proton jump, following photo-excitation of 7-AI dimer have been investigated by several groups [2] and is quite debatable. The transient absorption shows two component with a rise time at tautomer fluorescence, one prompt of 5 ps and the other a slower component of about 1 ns. These findings were interpreted in terms of a double minimum potential with a barrier of 500 to 700 cm^{-1} between dimer and tautomer [2]. Takeuchi and Tahara [1] have also reported the fluorescence dynamics of 7-AI, which consistently deny the existence of the intermediate species during the double proton transfer of the 7-AI dimer and supported the concerted mechanism in solution. In general, at room temperature a weak while at a low temperature of 77 K a strong emission corresponding to the dimer was observed.

Moreover, 7-AI is the chromophoric fraction of non-natural amino acid 7-azatryptophan, which makes it a good alternative of tryptophan for investigating proteins mechanism [5,6). The advantage of the 7-AI over the tryptophan is its optical features, which differentiate from tryptophan. These are the red-shifted absorption/fluorescence spectrum, higher fluorescence quantum yield, single exponential decay and longer lifetime.

ACCEPTED MANUSCRIPT

Further, the fluorescence of 7-AI is very sensitive to the solvent [7]. The absorption spectrum of 7-AI is formed by two overlapping electronic transitions from ??- ??* electronic state and assigned as ^1L_a and ^1L_b bands and behaves differently in solvents [8-10]. As an example, the ^1L_a electronic state interacts more strongly with the medium than ^1L_b state. Thus the Interaction between 7-AI and solvent molecules may be different in the photo-excited states than the ground state, which causes difference in charge distribution resulting variation in dipole moments. The excited state dipole moment is always important to evaluate the tunability range of the emission energy as a function of the polarity of the medium for a chromophore [11].

Various methods including electro-optic, Stark splitting, the electronic polarization of fluorescence, microwave conductivity, and electric dichroism have been used to calculate the dipole moment of large and small molecules and recently the nanoparticles/quantum dots. However, these methods are equipment sensitive therefore these methods are bound and restricted to facile molecules and quantum dots [12-16]. Among all the available methods, the most important especially for the investigation of the permanent electric dipole moments in the electronic ground and excited states is the solvatochromic methods given by Lippert–Mataga [17-19], Bilot-Kawski [20-22] and Bakhshiev [23]. The methods were also discussed by other groups [24-29] and used extensively for estimating the ground and excited states dipole moments [11,28,29]. These methods based on the variation of Stokes shifts with bulk solvent polarity described by dielectric constant (ε) and refractive index (η), and dimensionless solvent polarity parameter $E_T{}^N$.

We have used both the solvatochromic shifts methods and electric field modulation spectroscopic method, i.e., the Stark spectroscopy and have examined the charge separation, exciton dissociation, electron-hole recombination/separation and especially

ACCEPTED MANUSCRIPT

determined the change in the dipole moment between ground and photo-excited states of organic molecules, organometallic complexes and quantum dots [11,13-16,30,31].

Here, we have investigated the effect of microenvironment on the absorption and fluorescence of monomeric unit of 7-AI. The ground and photo-excited states dipole moments hence the change in the dipole moments were investigated separately in aprotic, protic and in all solvents using solvatochromic methods and DFT/DTDFT quantum chemical calculations. We have observed a significant charge separation in the photo-excited state, which critically depends on solvents used. Also, we have found that the change in fluorescence of 7-AI is nearly linear with the solvent polarity.

2. Theoretical Background

The ground and excited electronic single states dipole moments of 7-AI were calculated following solvatochromic shift methods, as follows:

2.1. Bakhshiev's formula [23]

$$\bar{\upsilon}_a - \bar{\upsilon}_f = S_1 F_1(\varepsilon, n) + Constant \qquad 1$$

where $\bar{\upsilon}_a$ and $\bar{\upsilon}_f$ are the wavenumbers (cm^{-1}) of the absorption and fluorescence maxima, respectively, ε and n are solvent dielectric constant and solvent refractive index, respectively. S_1 and F_1 are the slope of the linear equation and the bulk solvent polarity function, respectively, defined as follows:

$$S_1 = \frac{2(\mu_e - \mu_g)^2}{hca_0^3} \qquad 2$$

and

$$F_1(\varepsilon, \eta) = \frac{2n^2+1}{n^2+2}\left\{\frac{\varepsilon-1}{\varepsilon+2} - \frac{n^2-1}{n^2+2}\right\} \qquad 3$$

ACCEPTED MANUSCRIPT

where *h* is Planck's constant, *c* is the velocity of light in vacuum. μ_g and μ_e are the ground and excited singlet states dipole moments, respectively and a_0 is the Onsager cavity radius of 7-AI.

2.2. Bilot-Kawski formula [20-22,24]

$$\frac{\bar{\upsilon}_a + \bar{\upsilon}_f}{2} = S_2 F_2(\varepsilon, n) + Constant \quad (4)$$

where the meaning of the symbols are same as given in equations 1 and 2.

$$S_2 = \frac{2(\mu_e{}^2 - \mu_g{}^2)}{hca_0^3} \quad (5)$$

$$F_2(\varepsilon, \eta) = \frac{2n^2+1}{2(n^2+2)}\left\{\frac{\varepsilon-1}{\varepsilon+2} - \frac{n^2-1}{n^2+2}\right\} + \frac{3}{2}\left\{\frac{n^4-1}{(n^2+2)^2}\right\} \quad (6)$$

The parameters S_1 and S_2 can be estimated from the absorption and fluorescence band shift, as defined in the Eqs. 1 and 4. S_1 and S_2 are the slopes which can be determined using equations 1 and 4, respectively. Assuming the ground and excited states are parallel, the equations 2 and 5 can be expressed as.

$$\mu_g = \frac{S_2 - S_1}{2}\left\{\frac{hca_0^3}{2S_1}\right\}^{1/2} \quad (7)$$

$$\mu_e = \frac{S_1 + S_2}{2}\left\{\frac{hca_0^3}{2S_1}\right\}^{1/2} \quad (8)$$

and

$$\mu_e = \frac{|S_1 + S_2|}{|S_2 - S_1|}\mu_g \; if (S_2 > S_1) \quad (9)$$

The value of solute cavity radius, Onsager cavity radius was calculated from the molecular volume according to Suppan's equation [26,27]

$$a_0 = \left\{\frac{3M}{4\pi\delta N}\right\}^{1/3} \quad (10)$$

where δ the density of probe molecule, N is the Avagedro number and M is the molecular weight.

ACCEPTED MANUSCRIPT

2.3. Dimensionless microscopic solvent polarity parameter ($E_T{}^N$) [32]

The change in the dipole moment between ground and photo-excited states can be evaluated directly using the solvatochromic method, a dimensionless microscopic solvent polarity parameter $E_T{}^N$:

$$E_T{}^N = \frac{E_T(solvent)-30.7}{32.4} \quad (11)$$

and

$$E_T{}^N = \frac{28951}{\lambda_{max}(nm)} \quad (12)$$

where, λ_{max} is peak wavelength in the red region of the intramolecular charge transfer absorption of the pyridinium-N-phenolate betaine dye [33,34]. This further correlated with the Stokes shift as:

$$\bar{\upsilon}_a - \bar{\upsilon}_f = 11307.6\left[\frac{\Delta\mu^2 {a_D}^3}{{\Delta\mu_D}^2 {a_0}^3}\right] E_T{}^N + constant \quad (13)$$

were $\Delta\mu = \mu_e - \mu_g$ is the change in dipole moments of the molecules between ground (g) and photoexcted (e) states, $\Delta\mu_D$ is the change in the dipole moment of the Betaine dye (used as reference), a_D is Onsager radii of Betaine dye and $\boldsymbol{a_0}$ is Onsagar radii of 7-AI molecule. The calculated value of a_0 for 7-AI is 3.353 Å.

3. Experimental and computational details

7-AI was purchased from Sigma-Aldrich (purity ~ 99%) and used without further purifications. All the solvents were used of HPLC and spectroscopic grade. Absorption spectra were recorded with the UV/VIS/NIR spectrometer, Lambda-750, Perkin-Elmer. Fluorescence spectra were collected by Horiba Jobin Yvon Fluorolog-3 spectrofluorometer equipped with double-grating at excitation as well as at emission monochromators (1200 grooves/mm) and Hamamatsu R928P photomultiplier tube (PMT). The excitation source was

ACCEPTED MANUSCRIPT

a 450 Watt CW xenon lamp. Emission and excitation slits width were selected as required and data were collected using the origin software supplied by Horiba.

Ground-state geometry optimizations of 7-AI were carried out using DFT methods with the B3LYP functional (scheme 1). The excited-state calculations were performed using TDDFT methods with the B3LYP functional. All quantum calculations were performed with the Gaussian 09 program [35] in the gas phase. Throughout the ground-state optimizations and excited-state studies, the TZVP basis set was selected [36]. Ground and excited states dipole moments were calculated for the singlet ground and excited states.

4. Results and discussion

The absorption and fluorescence spectra of 7-AI have been measured at room temperature in various aprotic and protic solvents. Fig. 1(a,b) shows the absorption and fluorescence spectra of the monomeric unit of 7-AI in selected solvents. The absorption maximum lies between 287 and 290 nm. In aprotic solvents, 7-AI shows two well-resolved absorption bands; the higher energy band is assigned as ^1L_a and lower energy band as a ^1L_b band (Fig. 1a) [7-10,37,38]. The fluorescence maximum shifts from 343 to 390 nm while changing the solvent dielectric constant and polarity, as shown in Fig. 1b. The maximum Stokes shift varies from 5689 to 9081 cm^{-1} (Table 1). The large values of Stokes shift indicate a higher amount of interaction between solute and solvent in the excited state, pointing out that solute-solvent interaction in the excited state is different from that of the ground state, which may increase the excited state dipole moment. The correlation between the change in fluorescence maximum as a function of solvent dielectric constant or solvent polarity function, i.e., the plot between fluorescence maxima as a function of the solvent polarity is nearly linear (Fig. 2), demonstrating that the fluorescence of 7-AI can be tuned with the

ACCEPTED MANUSCRIPT

solvent polarity function, or the particular solvent can be determined by measuring the fluorescence of 7-AI. On the other hand, we could not observe such linear behavior between the solvent dielectric constant and fluorescence maxima, demonstrating that the increase in the Stokes shift with solvents predominantly due to the solvent polarity.

To ascertain charge separation in the photo-excited state and to estimate the dipole moment in the ground and photo-excited electronic states separately, the solvent polarity functions $F_1(\varepsilon, n)$ and $F_2(\varepsilon, n)$ have been calculated using equations 3 & 6 and are given in Table 1. Then the plots between spectral shifts ($\bar{\upsilon}_a - \bar{\upsilon}_f$) versus solvent polarity function $F_1(\varepsilon, n)$ and ($\bar{\upsilon}_a + \bar{\upsilon}_f$)/2 versus $F_2(\varepsilon, n)$, respectively were obtained separately for aprotic solvents (AS), protic solvents (PS) and considering all solvents (AS+PS). However, the plots are shown only for AS+PS in figures 3 and 4. The values of slopes obtained from the linear fitting curves of the Bakshiev correlation (S_1) and Bilot-Kawski corrections (S_2) are given in Table 2. Assuming the fact that ground and excited states dipole moments of the probe molecule are parallel, the values of μ_g and μ_e were calculated using eqs 7 and 8 and are surmised in Table 2. The values of the μ_g, μ_e and $\Delta\mu$ are 0.75, 3.21 and 2.46 D, respectively in aprotic solvents and increases slightly while considering all solvents (AS+PS). Especially, in protic solvents these values are nearly twice of that obtained in aprotic solvents (Table 2). It has been observed that the absolute value of the μ_g, μ_e and $\Delta\mu$ are different in the different set of solvents. The higher values of excited state dipole moment, hence the change in dipole moment ($\Delta\mu$) demonstrating the large charge separation in the photo-excited state resulting into the large Stokes shifted fluorescence. It has been reported that deviation from linearity occurs when excited and ground state dipole moments are antiparallel to each other. This only happens in case of large charge transfer character and noticeable for some of the aprotic polar solvents.

ACCEPTED MANUSCRIPT

Further, the change in the dipole moment between the two electronic states is evaluated using another solvatochromic method, a dimensionless microscopic solvent polarity parameter $E_T{}^N$, as described by eqs 11 and 13. The plot obtained between the solvent polarity parameter ($E_T{}^N$) versus Stokes shift ($\bar{\upsilon}_a - \bar{\upsilon}_f$) is nearly linear (Fig. 5). From the slope of the linear fit the change in dipole moment between the ground and photo-excited states is calculated using equation 13, considering the dye as reference. Typically the calculated value of $\Delta\mu$ is 1.84 D, which is smaller than the 2.68 D obtained using previous method. On the other hand, while considering the ^1L_b band the values becomes smaller confirming that the ^1L_a electronic state interacts more strongly with the medium than ^1L_b state [39].

The optimization of the ground and first excited singlet states together with calculations of ground and excited electronic states dipole moments of 7-AI were performed at the B3LYP/tzvp level, and the optimized structure is shown in Scheme 1. The calculated electronic transition energies and corresponding oscillator strength (*f*) of the low-lying singlet excited states are given in Table 3. The bright state (S_1) corresponds to the orbital transition of the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO) is HOMO to the LUMO. The ***CI*** value, the contribution of orbital transitions for other transition is comparatively small (Table 3). The slightly more bright state (S_2) is corresponding to the HOMO-1 to the LUMO. Typically, the S_4 state seems the brightest one and the CI value is maximum corresponds to the HOMO to the LUMO+1. The observed absorption and emission spectra obtained at B3LYP/tzvp level are given in Fig. 6 and demonstrating the vertical transition of 7-AI and are blue shifted as compared to the achieved in solvents. The calculated frontier molecular orbitals, HOMOs and LUMOs and the corresponding orbital energies of 7-AI are shown in Fig. 7, which show charge transfer

ACCEPTED MANUSCRIPT

character. Both ground and excited electronic states dipole moment were estimated and are given in Table 3. The theoretical calculations predict that the value of the excited state dipole moment is larger than the value of the ground state due to high polarity of the excited state, which is in agreement with the experimental results. It is worth mentioning here that examples are found for variation of dipole moments from the experimental results depending on the quantum chemical methods and basis sets used [9,10, 11,38,39,]. A variation in dipole moment is also noticed while going from gaseous to solvent phases, i.e., higher values of dipole moments are observed in the solvent as compared to the gas phase. This is due to the intramolecular interactions between the solvent molecules and the 7-AI. Further, the value of the dipole moment is found to increase in protic solvents. Thus, the observed results are equally important for understanding the large Stokes shifted fluorescence/tautomer fluorescence of 7-AI [1,30,31,37] and also reflecting that the CT character is favorable in 7-AI due to two functional groups attached nearby position of two different rings. Note that here we have not considered the dimer emission, as we have taken the monomeric unit of the solute.

The ground and excited state dipoles moments of 7-AI has also been examined using the Stark effect theory and found to be 1.59 and 2.30 D, respectively for S1 state, and these values are 1.59 and 4.60 D, respectively for the +280 cm^{-1} vibronic band at 286 nm [38]. The significant amount of permanent dipole moment and the variation in the dipole moment follow the similar trend as we observed experimentally and theoretically. It has been reported [38] that a significant mixing of π-π^* and n-π^*states leads to a major change in the magnitude of dipole moment. In another report [10] it has been concluded that all the four lowest-lying excited states namely ^1L_b π-π^*, ^1L_a π-π^*, n-π^* and π-σ^* crosses each other in one way or another, and hence significant state mixing occurs between them. The dipole

ACCEPTED MANUSCRIPT

moment of indole and 3-methylindole doped in polymer film have also been calculated using the Stark absorption spectroscopy [16]. The ground state dipole moment of indole and 3-methylindole is 1.42 D and 1.28 D and the change in the dipole moment is 5.6 and 7.2 D, respectively, and shows the restricted behavior of field induced alignment in the polymer films similar to that observed for 2- and 6-hydroxyquinolines [13,14]. Note that these values are not corrected by the internal field factor of polymer films.

5. Conclusions

The change in the dipole moment ($\Delta\mu$) between the ground and first excited electronic states, and separately the μ_g and μ_e of 7-AI have been calculated experimentally using solvatochromic shift method and theoretically using DFT/TDDFT calculations. The permanent electric dipole moment of 7-AI in the electronic excited state is significantly higher than the ground state, obtained both experimentally and theoretically. The difference between ground and excited state dipole moment obtained experimentally and theoretically may be due to the mixing of the lowest electronic excited states of 7-AI with one/or more higher excited states. The higher values of excited state dipole moment than the ground state dipole moments of 7-AI demonstrating the large charge separation in the excited state following photoexcitation, and also supporting the efficient proton transfer process of 7-AI. A nearly linear relation has been obtained between fluorescence and solvent polarity parameter, demonstrating the sensing ability of the 7-AI fluorescence with solvent polarity. Thus, the study is important to understand the charge transfer and proton transfer mechanism of DNA base pair.

ACCEPTED MANUSCRIPT

Acknowledgements

The work is financially supported by the Department of Atomic Energy (DAE), Board of Research in Nuclear Sciences (BRNS) (Grant No. 2012/37P/20/BRNS/765), Govt. of India. MSM thanks to Professor Keli Han, Dalian Institute of Chemical Physics, Chinese Academy of Science, China for providing the computational facility and lab members/project staff for their help.

References

[1] S. Takeuchi, T. Tahara, Chem. Phys. Lett., 347 (2001) 108, ibid, Proc. Natl. Acad. Sci. USA, 104 (2007) 5285-5290.

[2] W.M. Hetherington, R.H. Micheels, K.B. Eisenthal, Chem. Phys. Lett., 66 (1979) 230.

[3] K.C. Ingham, M.A. El-Bayoumi, J. Am. Chem. Soc., 96 (1974) 1674.

[4] X. Zhang, Y. Komoto , K. Sakota , N. Masayuki,T. Shinmyozu , S. Nanbu , H. Nakano, H. Sekiya, Chem. Phys. Lett. 443 (2007) 194–198.

[5] L. X. Q. Chen, J. W. Petrich, G. R. Fleming, A. Perico, Chem. Phys. Lett. 139 (1987) 55-61.

[6] B. Donzel, P. Gauduchon, Ph. Wahl, J. Am. Chem. Soc. 96 (1974) 801-808.

[7] A. V. Smirnov, D. S. English, R. L. Rich, J. Lane, L. Teyton, A. W. Schwabacher, S. Luo, R. W. Thornburg, and J. W. Petrich, J. Phys. Chem. B 101 (1997) 2758-2769.

[8] C. Kang, J. T. Yi, D. W. Pratt, Chem. Phys. Lett. 423 (2006) 7-12.

[9] Y. N. Svartsov, M. Schmitt, J. Chem. Phys. 128 (2008) 214310.

[10] S. Arulmozhiraja, M. L. Coote, J.Y. Hasegawa, J. Chem. Phys. 143 (2015) 204304.

[11] M.S. Mehata, A.K. Singh, R. K. Sinha, Methods Appl. Fluoresc. 4 (2016) 045004.

[12] J. Czekalla, Z. Chim. 15 (1961) 26, ibid, Z. Elektrochem. 64 (1960) 1221.

ACCEPTED MANUSCRIPT

[13] M. S. Mehata, T. Iimori, T. Yoshizawa, N. Ohta, J. Phys. Chem. A 110 (2006) 10985-10991.

[14] M. S. Mehata, K. Awasthi, T. Iimori, N. Ohta, J. Photochem. Photobiol. A 204 (2009) 39-45.

[15] M.S. Mehata, Sci. Rep. 5 (2015) 12056.

[16] E. Jalviste, N. Ohta, J. chem. Phys. 121 (2004) 4730-4739.

[17] E. Lippert, Z. Naturforsch. A 10 (1955) 541-545.

[18] E. Lippert, Z. Electrochem. 61 (1957) 962-975.

[19] N. Mataga, Y. Kaifu, M. Koizumi, Bull. Chem. Soc. Jpn. 29 (1956) 465-470.

[20] L. Bilot, A. Kawski, *Z. Naturforsch.* 17A (1962) 621–627.

[21] L. Bilot, A. Kawski, *Z. Naturforsch.* 18A (1963) 10–15.

[22] A. Kawski, P. Bojarski, *Spectrochem. Acta A* 82 (2011) 527-528.

[23] N.G. Bakhshiev, Opt. Spectrosc. 16 (1964) 821-832.

[24] A. Kawski, Z. Naturforsch. 57a (2002) 255-262, ibid, Naturwissenschaften 51 (1964) 82-83.

[25] A. Chamma, P.C.R. Viallet, Acad. Sci. Ser. C 270 (1970) 1901-1904.

[26] P. Suppan, C. Tsiamis, Spectrochim. Acta A 36 (1980) 971-974.

[27] P. Suppan, Chem. Phys. Lett. 94 (1983) 272-275.

[28] C.J. Talone, T.J. Gao, J.R. Lynch, R.M. Tanu, S.T. Deyrup, *Spectrochem. Acta A* 156 (2016) 138-142.

[29] B. Siddlingeshwar, S.M. Hanagodimath, *Spectrochem. Acta A* 72 (2009) 490-495.

[30] M.S. Mehata, C.S. Hsu, Y.P. Lee, N. Ohta, J. Phys. Chem. C 116 (2012) 14789−14795.

[31] M.S. Mehata, N. Ohta, Appl. Phys. Lett. 98 (2011) 181910.

[32] C. Reichardt, Chem. Rev. 94 (1994) 2319–2358.

ACCEPTED MANUSCRIPT

[33] S. Kumar, V.C. Rao, R.C. Rastogi, Spectrochim. Acta A 57 (2001) 41-47.

[34] M. Ravi, A. Samanta, T.P. Radhakrishan, J. Phys. Chem. 98 (1994) 9133-9136.

[35] M.J. Frisch, et.al. Gaussian 09, Revision C.01, Gaussian Inc, Wallingford CT, 2010.

[36] M.S. Mehata, Y. Yang, Z.J. Qu, J.S. Chen, F.J. Zhao, K.L. Han, RSC Advances 5 (2015) 34094–34099.

[37] K. Sakota, A. Hara, H. Sekiya, Phys. Chem. Chem. Phys. 6 (2004) 32-36

[38] J.W. Young, Z.D. Pozun, K.D. Jordan, D.W. Pratt, J. Phys. Chem. B 117 (2013) 15695-15700.

[39] L. S. Andres, M. Merchan, A. C. Borin, J. Stalring, Int. J. Quant. Chem. 84 (2001) 181-191.

[40] A. Frisch et al. Gaussian 09, Gaussian Inc. 2009.

[41] W. M. Haynes, CRC Handbook of Chemistry and Physics, 96th ed. Taylor & Francis, 2015.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 1. Physical parameters of solvents and spectral data of 7-AI in various solvents.

Type	**Solvent**	E_T^N	ε[a]	n[b]	F_1	F_2	$(\bar{\upsilon}_a - \bar{\upsilon}_f)$ **cm^{-1}**	$(\bar{\upsilon}_a + \bar{\upsilon}_f)/2$ **cm^{-1}**
AS	cyhex	0.006	2.0	1.4235	-0.00616	0.28432	5689	31999
	bz	0.111	2.3	1.5011	0.00992	0.34304	6313	31566
	tol	0.099	2.4	1.4941	0.03488	0.35100	6474	31485
	xyl	0.074	2.4	1.5018	0.0300	0.35353	6394	31525
	Et_2O	0.120	4.2	1.3524	0.3645	0.42219	6151	31647
	chl	0.259	4.7	1.4459	0.36177	0.48305	6591	31307
	ACN	0.460	35.7	1.3284	0.86301	0.65524	7695	30875
	DMSO	0.444	46.7	1.4793	0.84042	0.74418	7745	30610
PS	n-hex	0.559	12.5	1.4178	0.67753	0.62239	7235	30865
	2-PrOH	0.546	19.3	1.3776	0.77366	0.64369	7529	30718
	EtOH	0.654	24.9	1.3611	0.81482	0.6532	7382	30792
	MeOH	0.762	32.6	1.3284	0.85443	0.65095	7887	30539
	H_2O	1.0	78.4	1.3333	0.91263	0.68337	9081	30182

AS and PS stand for aprotic solvent and protic solvent, respectively. [a]*Ref. 40 and* [b] *Ref. 41.*

Solvent abbreviations as follows: Aprotic solvents: Cyclohexane (cyhex), benzene (bz), toluene (tol), xylene (xyl), diethyl ether (Et_2O), chloroform (chl), acetonitrile (ACN) and dimethylsulfoxide (DMSO).

Protic solvents: n-hexanol (n-hex), 2-propanol (2-PrOH) ethanol (EtOH), methanol (MeOH) and water (H_2O).

ACCEPTED MANUSCRIPT

Table 2: Ground state and the photo-excited state dipole moments of 7-AI.

a_0 (Å)	Solvent	Slope (cm^{-1})		Electronic States		Δμ (D)	
		S_1	S_2	μ_g (D)	μ_e (D)		
3.353	AS+PS	1905	-3205	0.91	3.59	2.68	This work
	AS	1603	-2584	0.75	3.21	2.46	
	PS	7132	-11509	1.59	6.78	5.19	
	AS+PS	2956			$\Delta\mu\ (E_T{}^N)$	1.84	
				1.59	2.30	0.71	Ref. 38
				1.59	4.60	3.01	Ref. 38
				1.63	2.1	0.46	Ref.9

AS and PS stand for aprotic and protic solvents, respectively.

Table 3: Calculated electronic transition energies and corresponding oscillator strengths of the low-lying singlet excited states and dipole moment of 7AI with the B3LYP method using TZVP basis set.

Electronic transition	Energy (nm/eV)	Oscillator strength	Contribution	CI %	μ_g (D)	μ_g (D)	Δμ (D)
$S_0 \rightarrow S_1$	268 (4.63)	0.0677	H →L	82.68	1.65	5.44	3.79
			H-1→L+1	1.81			
$S_0 \rightarrow S_2$	254 (4.89)	0.0745	H →L	7.52			
			H-1→L	69.13			
$S_0 \rightarrow S_3$	248 (5.01)	0.0030	H-2 →L	99.06			
$S_0 \rightarrow S_4$	204 (6.09)	0.3643	H-1→L	13.71			
			H-1 →L+1	2.24			
			H→L	2.68			
			H→L+1	70.81			
$S_0 \rightarrow S_5$	202 (6.15)	0.0001	H-2→L+1	99.47			

ACCEPTED MANUSCRIPT

H stands for highest occupied molecular orbital (HOMO) and L for lowest unoccupied molecular orbital (LUMO).

Captions:

Scheme 1. Ground state optimized structure of 7-AI with the B3LYP method using TZVP basis set.

Figure Captions:

Fig. 1. Normalized absorption (a) and fluorescence (b) spectra of 7-AI in different solvents, cyclohexane, acetonitrile, and water. The excitation wavelength was the peak absorption.

Fig. 2. Plot of fluorescence maxima versus microscopic solvent polarity parameter ${E_T}^N$.

Fig. 3. Bakhshiev correlation between the solvent spectral shifts and solvent polarity functions for all the solvents.

Fig. 4. Bilot-Kawski correlation between the solvent spectral shifts and solvent polarity functions for all the solvents.

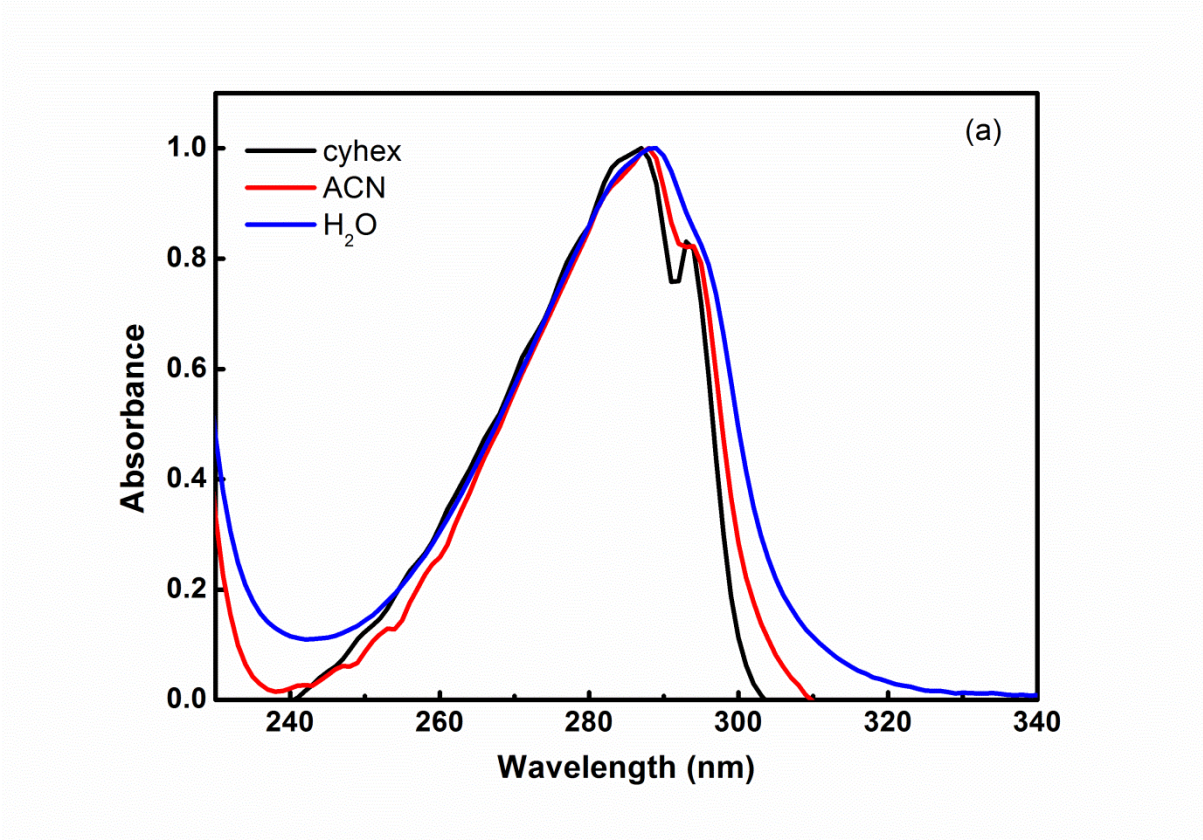
Fig. 5. Plot for Stokes shift versus microscopic solvent polarity parameter ${E_T}^N$.

Fig. 6. Calculated vertical transitions (absorption and emission) spectra of the low-lying singlet excited states of 7-AI with the B3LYP method using TZVP basis set.

Fig. 7. The calculated frontier molecular orbitals for the low-lying bright states of 7-AI with the B3LYP method using TZVP basis set.

ACCEPTED MANUSCRIPT

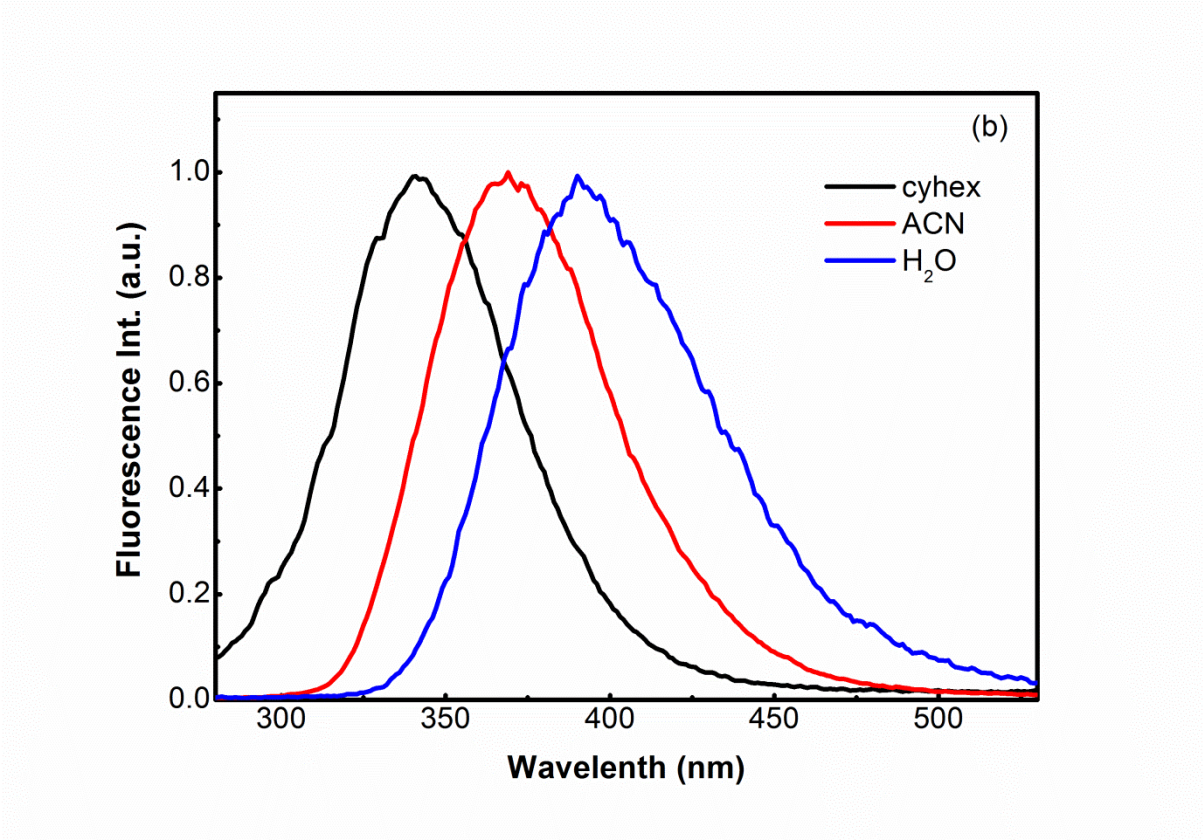
Figure 1a



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

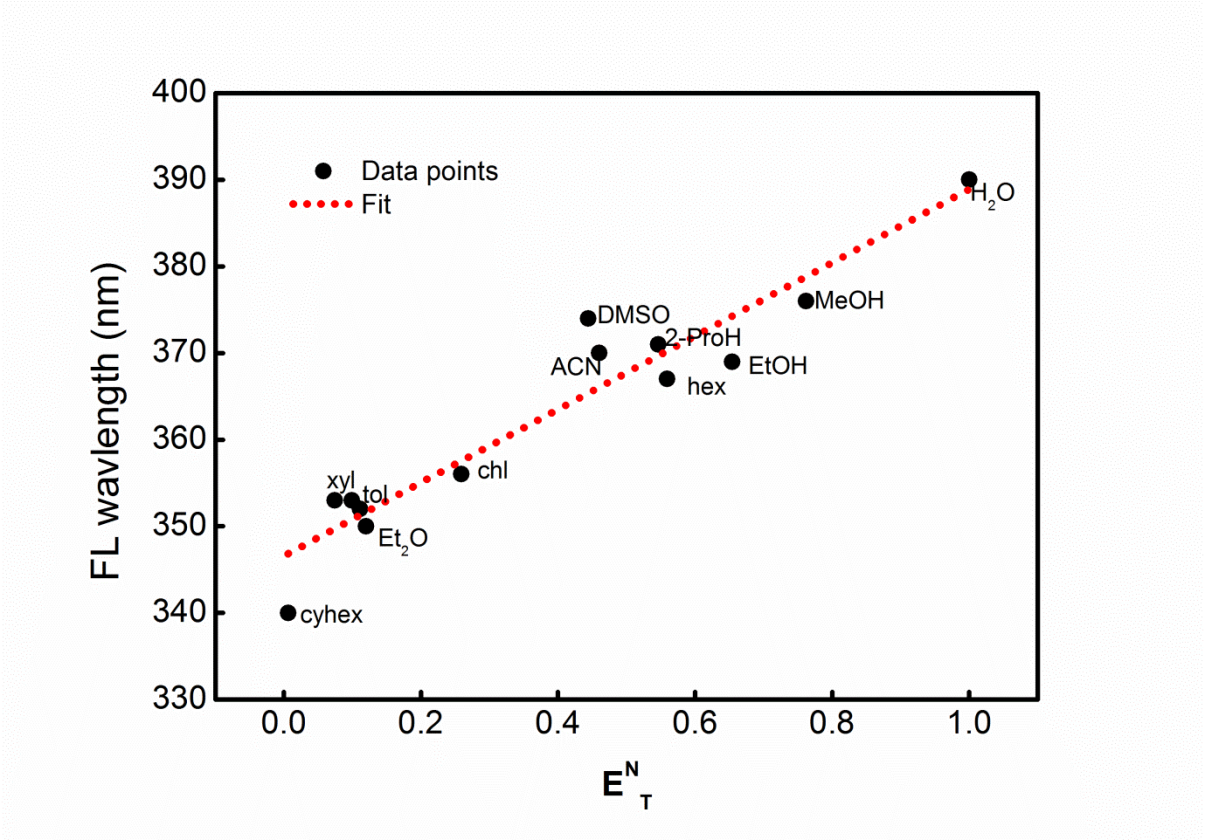
Figure 1b



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

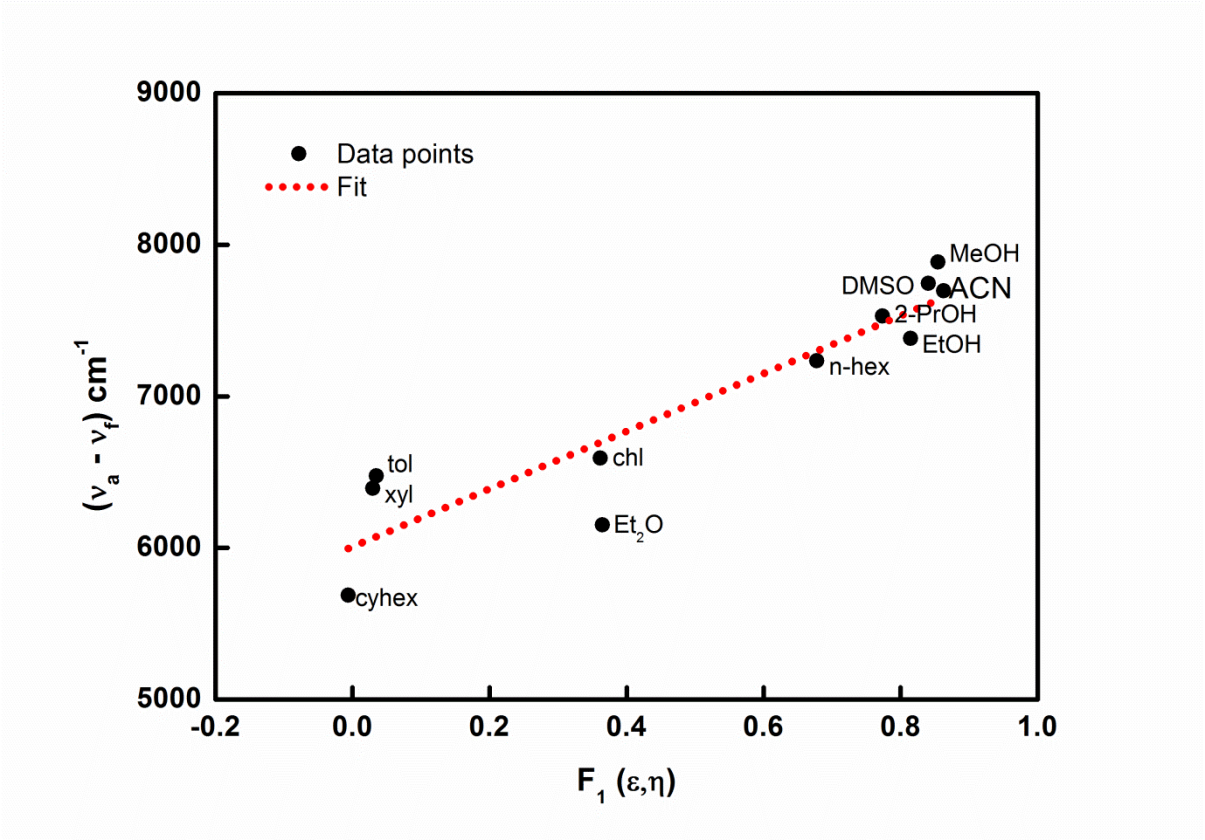
Figure 2



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

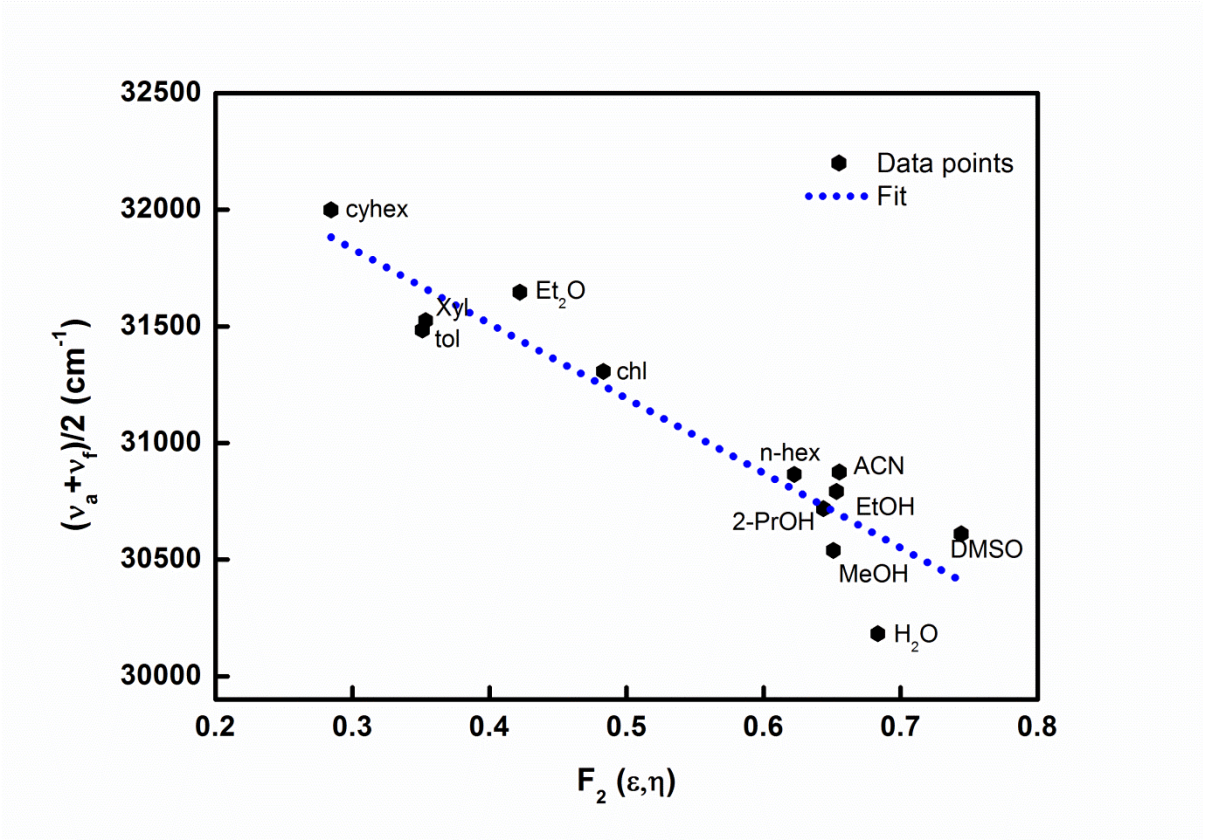
Figure 3



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

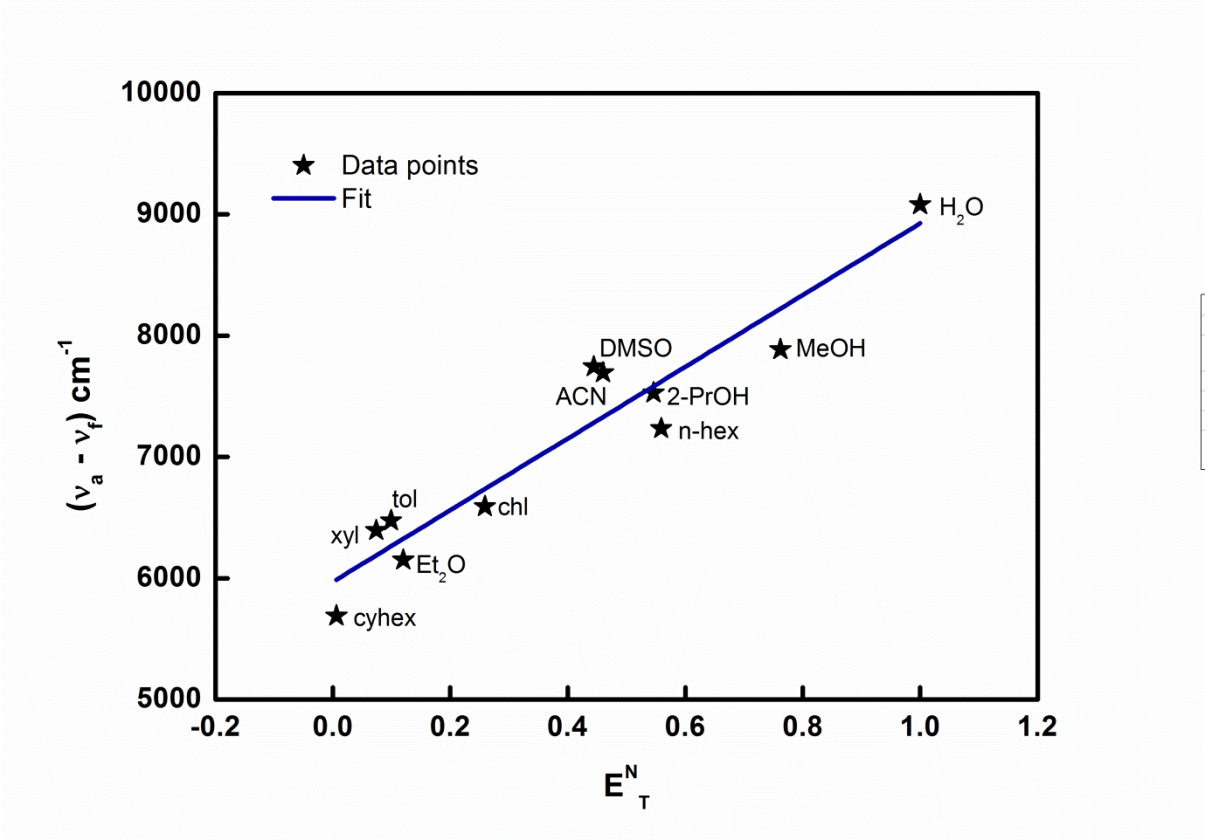
Figure 4



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

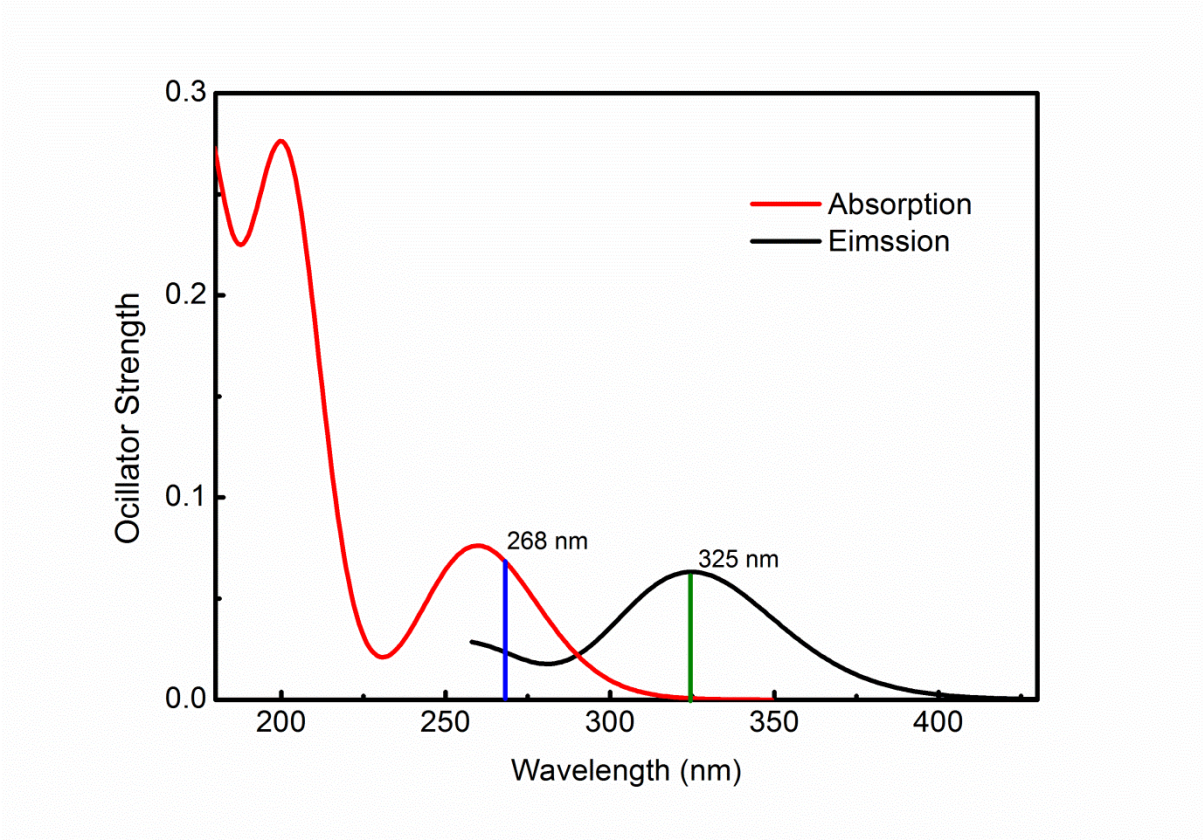
Figure 5



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

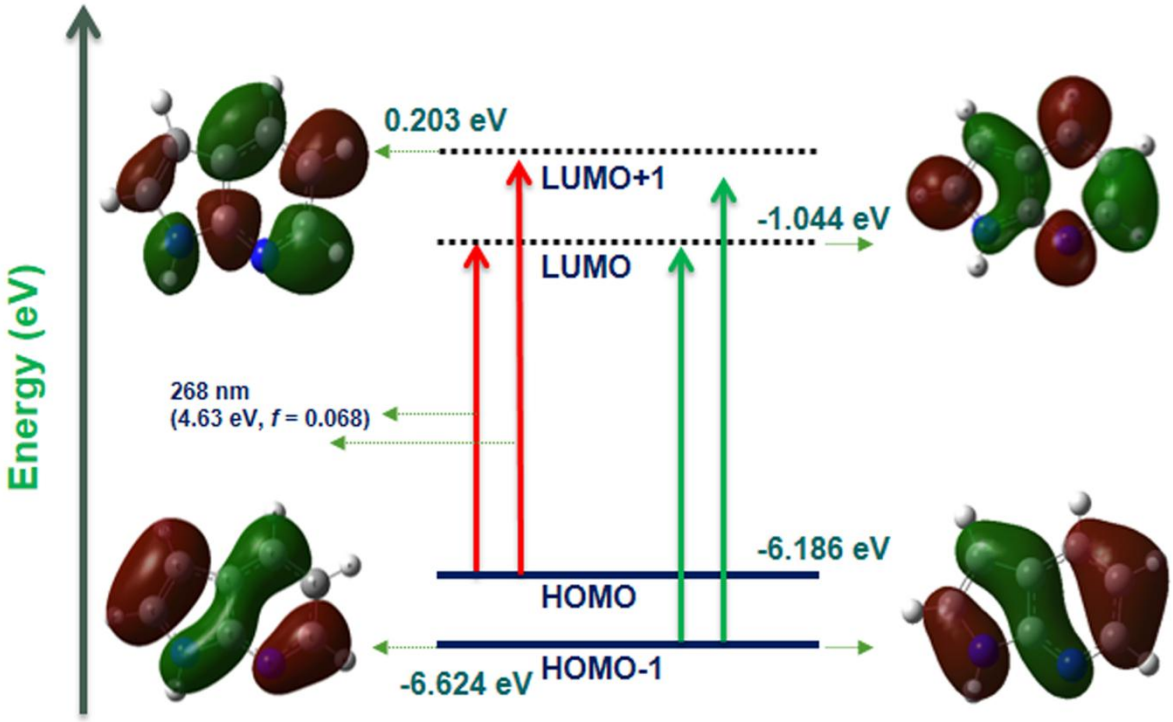
Figure 6



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

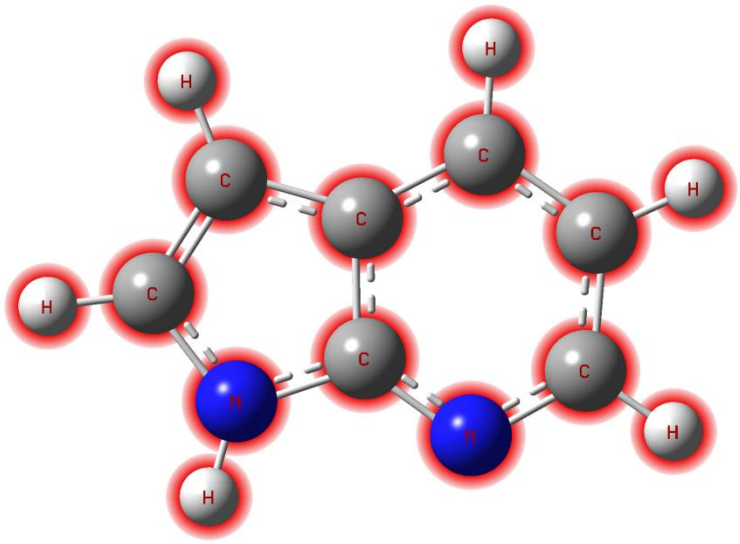
Figure 7



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Scheme 1



ACCEPTED

ACCEPTED MANUSCRIPT

Highlights:

- The vertical transitions, bright states and HOMO-LUMO of 7-AI were obtained from DFT/TDDFT.
- Ground and excited states dipole moments have been estimated experimentally and theoretically.
- Significant charge separation has been observed in 7-AI following photoexcitation.
- The relation of fluorescence and solvent polarity indicates the sensing ability of 7-AI to solvent.

ACCEPTED MANUSCRIPT

Journal of Information Security and Applications 000 (2017) 1–15

Contents lists available at ScienceDirect

Journal of Information Security and Applications

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

Localized & self adaptive audio watermarking algorithm in the wavelet domain

Arashdeep Kaur[a,*], Malay Kishore Dutta[a], K.M. Soni[a], Nidhi Taneja[b]

[a] Amity School of Engineering & Technology, Amity University Uttar Pradesh, Noida, India
[b] Delhi Technological University, Delhi, India

ARTICLE INFO

Keywords:
Adaptive audio watermarking
Wavelet decomposition
Payload estimation

ABSTRACT

This paper presents an adaptive audio watermarking algorithm in the wavelet domain to optimize the payload under the perceptual transparency constraints of audio signal by strategically using some of its local features. Unlike existing algorithms, the watermark payload in this approach is made adaptive based on the nature of the audio signal. This localized feature based approach to determine the payload addresses the issue of over-loading and under-loading the audio signals with watermark data making the payload optimized for each individual audio host signal. Some audio features are strategically extracted and the most discriminatory features are selected using Principal Component analysis (PCA) approach. A mathematical model is designed using selected audio features like energy, zero cross mean and short time energy to evaluate the degree of embedding under perceptual transparency. It is used to estimate the number of watermarking bits to be inserted for a particular audio signal which makes the approach adaptive in nature optimizing the watermarking payload. At the embedding stage, watermark is embedded in the host audio signal in the third level detailed coefficient of wavelet domain which strikes a balance between the contradicting design parameters of perceptual transparency, robustness and optimized payload. Watermark extraction in this paper is blind with good robustness to signal processing attacks. Experimental results validate that the proposed adaptive algorithm provide good imperceptibility with good robustness against signal processing attacks at adjustable payload for different types of audio signals. Comparative analysis indicates that this proposed adaptive algorithm has better performance in terms of imperceptibility and robustness in comparison to uniform watermarking algorithm.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Watermarking of audio signals have been in existence since last few years for secure transmission of audio signals over the network addressing copyright issues, piracy issues, ownership issues, fingerprinting and military applications etc. There are wide range of audio watermarking algorithms that are available in the existing literature for varied applications. Audio watermarking can be done in time domain, frequency domain, wavelet domain, cepstrum domain etc. Audio watermarking algorithms are basically divided into two categories: blind and non-blind. If original signal or watermark is required at the receiver side, then the watermarking algorithm is said to be non-blind, else it is blind algorithm. It is very challenging to design a blind audio watermarking algorithm because detection of efficient watermark at the receiver end, after application of various signal processing attacks is a crucial task. The embedding algorithm should be effective enough such that the original signal and the watermarked signal are perceptually similar. The watermarked signal must withstand serious signal processing attacks in order to extract the exact watermark with good accuracy. Imperceptibility, robustness and payload are the three contradictory design requirements of an audio watermarking algorithm [1]. In order to design an efficient audio watermarking algorithm, the three contradictory design parameters of the watermarking algorithm should be optimized. There are a large number of existing algorithms which provide a good watermarking solution at a high payload with good robustness against signal processing attacks maintaining perceptual constraints [2–6]. However, these algorithms may be suitable for some particular set of audio signals because property of audio signals is not taken into consideration while designing the embedding and extraction algorithm. In order to provide a solution to this problem, many researchers have proposed adaptive audio watermarking algorithm in which embedding strength of the algorithm is decided by using some of the audio features. The embedding strength here represents quantization pa-

* Corresponding author.
E-mail addresses: akaur@amity.edu, arashgulati@gmail.com (A. Kaur), mkdutta@amity.edu (M.K. Dutta), kmsoni@amity.edu (K.M. Soni), nidhi.iitr@gmail.com (N. Taneja).

http://dx.doi.org/10.1016/j.jisa.2016.12.003
2214-2126/© 2017 Elsevier Ltd. All rights reserved.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

rameter or some alpha factor etc. Li et al. has presented an adaptive audio watermarking algorithm in wavelet domain based on SNR to determine watermark embedding intensity by using scaling parameter,α [7]. Pooyan et al. [8] proposed a robust method of audio watermarking in wavelet domain which determines the quantization and embedding strength adaptively according to the characteristics of human auditory system (HAS). Wang et al. presented an adaptive audio watermarking algorithm in discrete multi wavelet transformation based on the energy relationship of two consecutive frames in multi wavelet domain [9]. Dahui Li presented an adaptive audio watermarking algorithm in DCT and DWT domain. This algorithm also chooses the quantization step according to the masking properties of HAS [10]. Dutta et al. proposed an efficient watermarking algorithm which embeds watermark data adaptively in SVD and DWT domain. In [11] high energy peaks are used to evaluate degree of embedding for each audio frame. Youssef presented a novel hybrid fuzzy self-adaptive digital audio watermarking scheme (HFSA-A W) in DWT which is based on local audio features [12]. Fuzzy c-means clustering is used in [12] for segmentation of audio features related to rhythm, timbre and harmony to estimate the strength of a frame for each sub-band and to ensure that the embedded watermark in the original audio is self-adaptive. But the method is non-blind because it requires an original signal at the receiver end for watermark extraction. Peng et al. presented an adaptive blind audio watermarking algorithm in wavelet domain based on local audio feature and support vector regression (SVR) [13]. In their paper, frame energy and maximal peaks of its all sub-bands are extracted as local features and SVR is used to model the relationship between the local features and the embedding strength of the audio frame in order to adaptively control the embedding strength of the audio frame.

It can be seen from the state-of-art that there exists various adaptive algorithms addressing the issue of robustness and imperceptibility in audio watermarking. In these algorithms, watermark is made inaudible by using different embedding intensities in different audio frames that depends on some features of that particular audio frame. Robustness is another design parameter of audio watermarking which is optimized with good perceptual transparency. Watermark payload capacity of the signal is not being considered in these methods. Different audio signals may have different capacity of carrying watermark data. For example, if an audio signal has low energy or less noise but carrying more watermarking bits than its capacity for fixed payload watermarking algorithm, it will lead to overloading of audio signal with watermark data. The changes due to insertion of watermark may be audible to human ear or it may also affect the robustness of the algorithm. In other case, if a high energy audio signal carry less number of bits than its capacity, this will lead to under loading of audio signal with watermark data. In this case robustness or imperceptibility will not be disturbed or even may be improved. Thus, there is an unused room available which should be utilized for hiding watermarking bits. In order to overcome the issue of overloading and under loading of audio signal with watermark data, an algorithm is proposed in this paper by considering the features of audio signal such that payload capacity directly depends on the signal leading to an adaptive nature of watermarking algorithm. This paper presents an adaptive audio watermarking algorithm in wavelet domain with varied payload for different audio signals. This proposed algorithm can estimate the payload for the audio signal by using local audio features like energy, zero cross mean etc.

The main contribution of this paper is an adaptive audio watermarking algorithm that can estimate the watermark payload by using local audio features from the host audio signal. This will optimize the payload that to an acceptable limit for an audio signal under perceptual transparency constraints. Unlike various existing algorithms that involves uniform payload, this adaptive algorithm will help in resolving the issue of over-loading and under-loading the audio signal with the watermark data. This will ensure that the most optimized payload is employed for each host audio signal depending on its local properties which makes the algorithm adaptive in nature. To address the design requirements like imperceptibility, robustness and adaptive payload which are mutually contradictory, the watermarking method is strategically chosen based on a wavelet decomposition method.

The rest of the paper is organized as follows: Section 2 describes the challenges faced while designing an adaptive audio watermarking algorithm, detailed methodology of the proposed algorithm is given in Section 3, experimental results are discussed in Section 4 and finally Section 5 concludes the paper and proposes a future prospect.

2. Adaptive audio watermarking: challenges

In any watermarking method, payload may affect both imperceptibility and robustness of the algorithm. Generally, more payload has less robustness and lower imperceptibility, whereas watermarking algorithms with less payload rates can retain watermark even under signal processing attacks exhibiting good imperceptibility and robustness of the algorithm. In general, existing audio watermarking techniques have fixed watermarking payload for any type of signal without considering the nature of the signal. This can cause the problem of under-loading and over-loading the audio signal. So, due to different natures of the audio signals, the payload capacity should differ from signal to signal. Adjusting the payload or embedding intensity on the basis of audio signal for an algorithm is considered as an adaptive audio watermarking. In this paper, an adaptive audio watermarking algorithm is proposed in which payload is estimated from local features of the audio signal. It can avoid the issue of over-loading or under-loading of audio signal keeping the payload optimized under perceptual transparency conditions and good robustness toward signal processing attacks. In an adaptive audio watermarking algorithm wherein payload of an algorithm varies from signal to signal, many challenges have been faced. Some of the challenges faced while designing adaptive algorithm are listed as below:

1. The major challenge while designing an adaptive audio watermarking is the selection of audio features which may impact the payload of the signal such that number of bits to be embedded may depend on the localized audio features.
2. The design of mathematical model to evaluate the payload of the signal from its inherent features is considered as this will serve as root for the adaptive algorithm to decide the payload based on the local property of the audio signal.
3. The effective design of embedding algorithm is one mandatory requirement for adaptive audio watermarking such that variable payload can be embedded in different types of audio signals without affecting the perceptual transparency and the robustness.
4. Evaluating the payload of the watermarked signal at the receiver side is also a challenging task needing attention. The embedding is done in such a way that the features used for evaluating the payload should not be disturbed by hiding the watermark.

Therefore, there is a need to design an adaptive audio watermarking algorithm maintaining good robustness under perceptual constraints with varying payload for different types of audio signals.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

3. Proposed methodology

This paper presents a localized audio feature based adaptive audio watermarking algorithm in wavelet domain having good robustness against signal processing attacks without disturbing the audible quality of the watermarked signal at different payload rates for different types of signals. In this paper, payload for an audio signal is found by evaluating the different features of the framed audio signal. This section describes the detailed methodology used for design and implementation of the proposed algorithm. This section also discusses the significance of adaptively estimating the payload of an audio signal by using the localized features. Further embedding procedure and extraction procedure for the proposed watermarking algorithm are also discussed in this section.

3.1. Justification for use of audio features to estimate payload

Audio signal differs from others in terms of wavelength, amplitude, frequency etc. Audio signals unveiling different properties should be treated differently to optimize the watermark embedding payload capacity in the signals. There are three design parameters of an audio watermarking algorithm namely imperceptibility, robustness and payload [1]. Imperceptibility means that there is no distortion in the audible quality of the signal after watermark embedding. Robustness refers to the ability of the algorithm to resist watermark under signal processing attacks. Payload or embedding capacity of an audio watermarking algorithm is defined as the number of watermarking bits which can be embedded reliably per unit of time in an audio signal. The payload of an audio watermarking algorithm can be evaluated by using Eq. (1).

$$Payload\ capacity = \frac{Number\ of\ watermarking\ bits}{Length\ of\ Audio\ Signal\ in\ seconds}\ bps \quad (1)$$

The payload of an audio signal is measured in bits per second (bps). Any changes made in the audio signal due to embedding of watermarking bits, should not disturb the audible quality of the signal. Imperceptibility can be affected directly by the number of watermarking bits per second in the signals. More the number of watermarking bits, more disturbance will be created in the watermarked signal and lesser the number of watermarking bits, a good quality watermarked signal can be generated. Similarly, robustness also depends on the number of bits the watermarked signal is carrying. Thus, watermark payload of an audio watermarking algorithm is a crucial design parameter to be taken formerly under consideration for optimized performance. While designing a watermarking algorithm, the payload of an algorithm depends on the method of embedding and sometimes on the region of embedding.

Various works have been reported in the existing literature aiming to find an optimum balance between the contradictory design requirements. Hwai et al. [14] have given an audio watermarking algorithm with payload of 473 bps within perceptual constraints. Their method have shown good robustness to all the attacks except compression. Fallahpour et al. proposed a FFT coefficients based 2000–6000 bps algorithm but this algorithm was not evaluated for any attack other than Mp3 compression [15]. It can be seen from the experimental results provided in the papers that different audio signal behaves differently under similar attack situation at stable payload. Huang et al. have given an optimized embedding algorithm based on SNR. In their paper, the payload has been optimized between 500 bps and 800 bps, but still the robustness against signal processing attacks is still a challenge.

In the existing works, the type or nature of audio signal is not taken into consideration while designing an audio watermarking algorithm which may have unbalanced impact on the payload of the watermarking. In general existing audio watermarking algorithms have fixed watermark payload for any type of audio signals without considering its nature and property. It can cause signal to carry more number of watermark bits than it can actually carry within transparency limits and can also cause some signal to carry lesser number of bits than its capacity to maintain transparency. For example, if an audio signal is having low energy and is overloaded with watermarking bits this may directly affect the imperceptibility and the robustness. If an audio signal has high energy and it carries lesser number of bits than its capacity leads to under-loading of watermark data in an audio signal, then the unutilized space available to hide the data can be utilized.

Hence a fixed payload audio watermarking algorithm in which signal inherent property is not taken into consideration to decide the watermarking payload may lead to overloading and under loading effect of payload within perceptual transparency limits and may lead to non-optimized robustness. Keeping in mind this issue, this paper proposes an adaptive audio watermarking algorithm which initially analyzes the audio signal to evaluate optimized payload under perceptual transparency constraints before watermark embedding.

The payload for each signal to be watermarked is decided individually based upon their local properties and watermark embedding method which may avoid the problem of over-loading and under-loading of the watermark in the audio signal. This makes the watermarking algorithm signal dependent for payload estimation optimizing the design requirements. Audio signals are analyzed in this paper by extracting the local features from them. Audio features which have been extracted in this paper are listed in Table 1 with its definition.

Table 1
Description of audio features.

S.No.	Audio feature	Description
1	Pitch	Measure of fundamental frequency of audio signal
2	Formants	Range of frequencies in which there is an absolute or relative maximum in the sound spectrum
3	Zero crossing rate	Number of times audio waveform crosses zero axis
4	Energy	Energy of the audio signal
5	Volume	Perception of energy used in voice production
6	Silence ratio	Percentage of low energy audio frames
7	Centroid	It is the "centre of mass" of audio spectrum
8	Roll off	Frequency below which 85% of the magnitude distribution is concentrated.
9	Flux	Rate of change in power spectrum of a signal
10	Short time energy	Energy of an audio frame
11	Skewness	Rate of change of waveform w.r.t time

These audio features have been chosen for the experimentation as these may have direct or indirect impact on the payload of the audio signal with relevance to the watermarking design parameters of perceptual transparency and robustness. For example, if centroid of the audio frame is large, it implies that the frame has high spike and then that particular frame can have high embedding capacity as embedding the watermark data in the high peaks will not be audible to human ear according to the pre-masking and post masking rules. Thus maintaining good perceptual transparency and robustness at high payload. This can help in resolving the problem of overloading and under-loading the audio signal with an optimized payload under the design limits.

According to psychoacoustic model, the sensitivity of the human ear to quantization noise in a given sub band is directly proportional to the level of audio energy in that sub band [16]. Hence, the payload of an audio signal will have relation with the energy of audio signal. To justify the relation between the energy of the audio signal and the SNR of an audio signal with some particular embedding method, we conducted some experiments. In Fig. 1, we

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

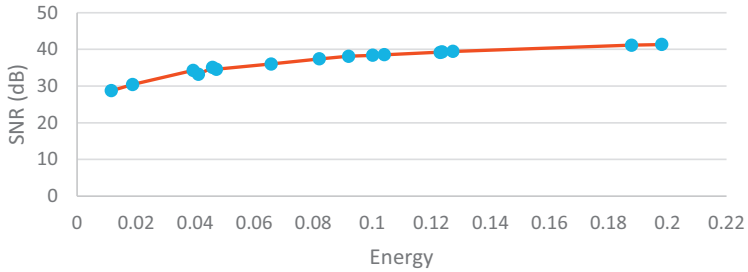


Fig. 1. Plot depicting the relation between SNR and Energy of an Audio Signal.

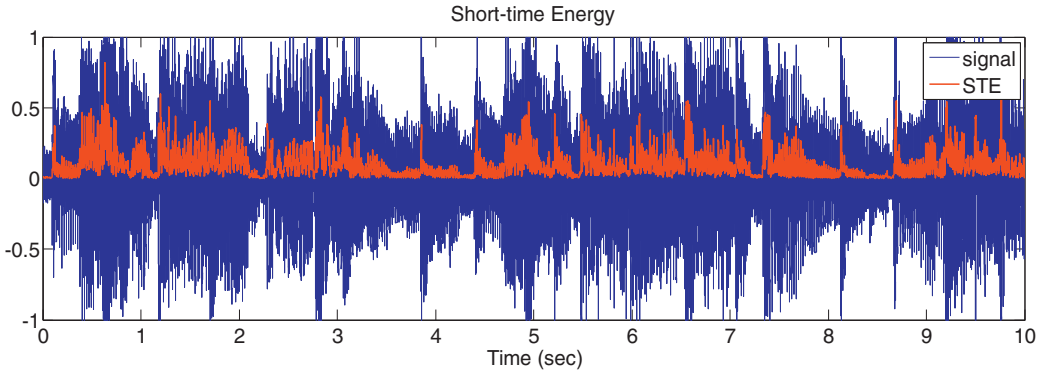


Fig. 2. Representation of short time energy (in red) of Blues audio. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

can observe that if an audio signal exhibits high energy then it can achieve high SNR at fixed payload rates. Thus, energy of an audio signal is used as a basic feature in this work to select other features which may impact the payload of the signal. The SNR of an audio signal has been evaluated using Eq. (2).

$$SNR\ (\mathrm{So}, \mathrm{Sw}) = 10log_{10}\frac{\sum_{i=1}^{L} \mathrm{S_o}^2\ (i)}{\sum_{i=1}^{L} [\mathrm{S_o}(i) - \mathrm{S_W}(i)]^2}\ dB \tag{2}$$

where S_o and S_w are the original and watermarked audio signals and L is the length of the audio signal.

The plot given in Fig. 1 represents the energy of an audio signal along X-axis and SNR of the watermarked audio signal generated by using the embedding method proposed in this paper at 640 bps payload along Y-axis. The plot has been obtained by using some of the audio signals from the database represented using the blue dots. This plot depicts the direct relation between the energy of the audio signal and the SNR by using the proposed embedding method. Thus, using the observations from the experimentation, energy is found to be correlated with the SNR if embedding is done using the proposed embedding method. Hence, energy of an audio signal is used as a base feature to estimate payload of the audio signal for the proposed embedding under perceptual constraints.

3.2. Adaptive algorithm

Audio feature selection in this paper is done to find the suitable audio features which may impact the watermarking payload. The main concentration in this paper is to establish the mathematical model to evaluate the audio payload from its local properties. Some features have been selected from the pool of large number of features extracted from different audio signals as given in the previous section. The feature selection in this paper has been done by using PCA (principal component analysis) algorithm as a classifier. Features of all the frames of the audio signal are evaluated and PCA is then applied in various iterations on the feature dataset to find the features which have similarity index closer to energy of the signal. It is then found that two audio features namely energy and short time energy are closely related to each other. Hence, if SNR of an audio signal is directly dependent on the energy of the audio signal then the features which are closely related to energy of the signal may impact the SNR of the signal. Thus, energy and short time energy of the framed signal are considered as discriminatory features that have direct relevance to acceptable payload of an audio signal.

Short time energy is defined as the energy of short audio segment [17]. Short time energy is the simple and effective classifying parameter used for finding nature of audio segments [18]. Relatively, high short time energy is an indication of high peaks in the short segments of audio signal and vice-versa, refer Figs. 2 and 3. Also, Zero Crossing Rate (ZCR) is used in this paper along with energy and short time energy to evaluate payload. ZCR counts the number of times spectrum crosses zero in a signal. Fig. 4(a) represents the waveform plot of blues audio signal and Fig. 4(b) gives the number of times a signal crosses zero for blues audio. These figures depict that high amplitude of a signal at any particular instance of time have less ZCR whereas low amplitude values have high ZCR.

It was found empirically after analyzing all audio signals of different natures in a database that a high energy signal have less ZCR as compared to low energy signal. It can be seen from Fig. 2, the audio signal having high amplitude value at a particular instance of time have high energy and vice-versa. Figs. 5 and 6 represent that the audio signal that have high amplitude at particular instance of time also have less zero crossing rate (ZCR) and vice-versa. Therefore, amplitude is directly proportional to energy and indirectly to ZCR. Also, we can say that the energy is indirectly proportional to ZCR. Hence, ZCR of a signal can be inversely proportional to payload of the signal, if energy is directly proportional to payload of a

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

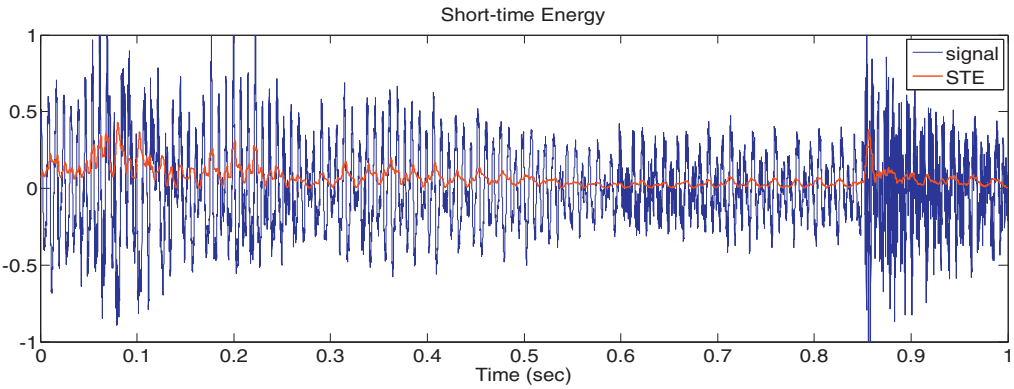


Fig. 3. Short time energy (in red) of Blues audio for 1 s. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

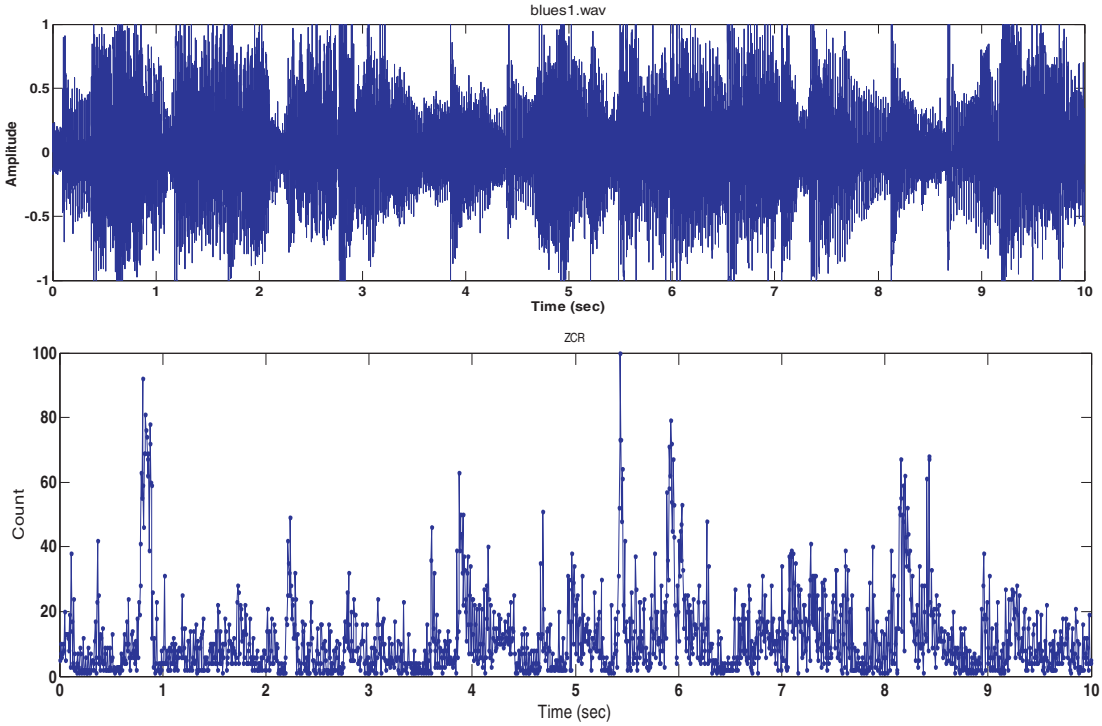


Fig. 4. (a) Blues audio signal. (b) Representation of zero cross mean of Blues audio.

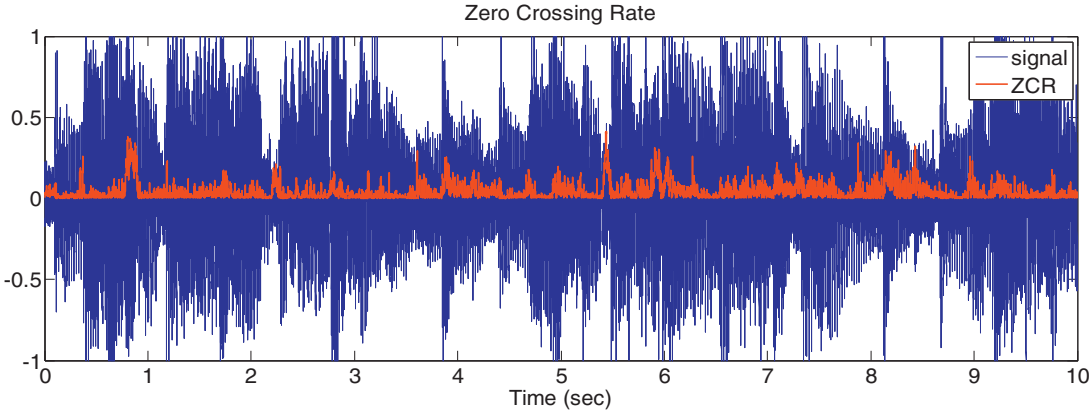


Fig. 5. Plot depicting ZCR (in red) for Blues audio. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

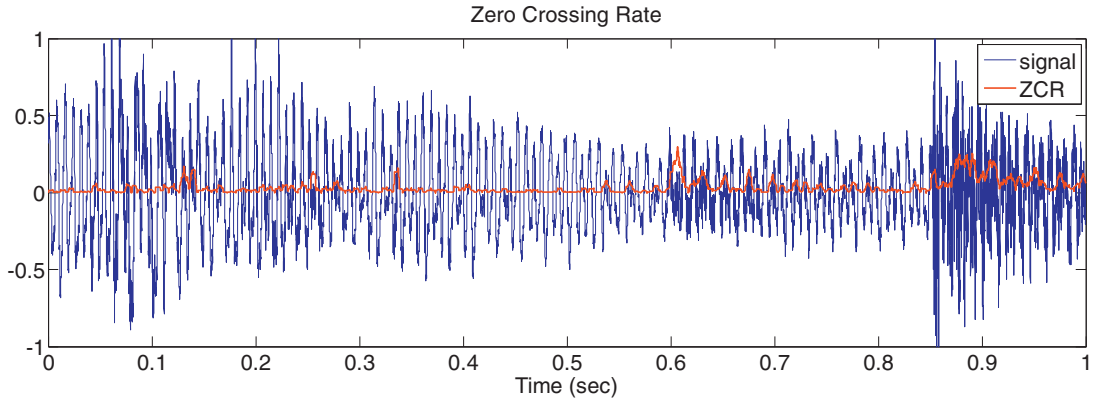


Fig. 6. ZCR (in red) of Blues audio for 1 s. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

signal. Thus, short time energy, energy and ZCR are the three audio features that are used for payload estimation in this paper.

Energy of an audio signal can also be used to detect whether the signal contains good opportunity for data embedding without affecting the perceptual transparency or it has low peaks to accommodate less number of watermarking bits keeping perceptual transparency at an acceptable rate. The energy of an audio signal, y, can be defined as below [19]:

$$E_n = \sum_{n=0}^{N-1} y^2(n) \tag{3}$$

where, y(n) denotes the discrete samples of signals at regular intervals. The short time energy of the above signal can be determined from the following expression:

$$STE_n = \sum_{n=m-M+1}^{m} [y(n)w(m-n)]^2 \tag{4}$$

where, w(m-n) is the window, m is the sample that the analysis window is centered on, and M is the window length. Here, high energy values are considered to be representing high peaks and vice versa. High peaks in audio signal gives good indication of availability of high capacity for data hiding within perceptual constraints.

On the basis of above discussion, it is found that energy and short time energy of the audio signal can be considered to be directly dependent on payload of the signal. To evaluate the payload of an audio signal, it becomes a mandatory requirement to find the strength of the audio signal. In this paper, audio strength of the signal has been evaluated by using energy, short time energy and ZCR as given in Eq. (5). The empirical formula based on the above mentioned logic to evaluate the audio strength is given as follows:

$$Audio_Strength = \frac{(Short\ Time\ Energy * x_1 + Energy * x_2)}{mod(\alpha,\ ZCR) * 10} \tag{5}$$

Where, mod function gives the remainder and α *is an adjustment parameter*,

$$x_1 = mean\ (all_{peaks} > 0.5)$$

$$x_2 = \frac{\int_0^{\pi} \frac{\sin t}{t} dt - \frac{\pi}{2}}{\pi}$$

Here, higher x_1 can be chosen for audio signals that have more number of higher peaks to embed more number of watermarking bits effectively in a robust manner without compromising imperceptibility according to law of large numbers [20]. The value of x_2 is found using the Gibb's phenomenon. A sine wave (sinusoidal function) can be considered as a basic building block which can be used for signal analysis. [21]. One steady sine wave can represent exactly one frequency present in a signal. This factor is multiplied with the energy of the signal to find the optimal audio strength. Higher the strength of the audio, more the payload a signal can carry without disturbing the audible quality of the signal.

Table 2
Multiplication factor for audio strength.

Range					Mul_Factor	Sign
0	<	Audio_strength	<	0.03	0.9	–
0.03	<=	Audio_strength	<	0.04	0.8	–
0.04	<=	Audio_strength	<	0.05	0.7	–
0.05	<=	Audio_strength	<	0.06	0.6	–
0.06	<=	Audio_strength	<	0.07	0.5	–
0.07	<=	Audio_strength	<	0.08	0.4	–
0.08	<=	Audio_strength	<	0.09	0.3	–
0.09	<=	Audio_strength	<	0.1	0.2	–
0.1	<=	Audio_strength	<=	0.15	0.1	–
0.15	<	Audio_strength	<	0.2	0.1	+
0.2	<=	Audio_strength	<	0.3	0.2	+
0.3	<=	Audio_strength	<	0.4	0.3	+
0.4	<=	Audio_strength	<	0.5	0.4	+
0.5	<=	Audio_strength	<	0.6	0.5	+
0.6	<=	Audio_strength	<	0.7	0.6	+
0.7	<=	Audio_strength	<	0.8	0.7	+
0.8	<=	Audio_strength	<	0.9	0.8	+
0.9	<=	Audio_strength	<	1	0.9	+

The strength of audio signal is used to find a multiplication factor which is then multiplied with the fixed payload of the watermarking algorithm to find the acceptable payload depending on the audio signal meeting the perceptual constraints with good robustness against signal processing attacks. A comprehensive experimentation done on large set of audio signals of various quality and nature indicates that audio strength can be related with the watermark payload. It was then found that the signal having audio strength greater than 0.15 can accommodate more payload and can increase the efficiency accordingly. With the exhaustive testing, multiplication factors are found on the basis of audio strength and are indicated using Table 2.

The optimized payload, number of watermarking bits, in this paper, is found by using the following equation:

$$watbits\ (sign) = watbits * Mul_Factor \tag{6}$$

Where, sign and Mul_Factor can be obtained from Table 2. Thus, Eq. (6) gives the optimum number of bits per second that can be embedded efficiently in a particular audio signal. This can help in resolving the problem of under loading and overloading of an audio signal. The value of α is chosen to be 0.01 for the experimenta-

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

tion in this paper. Watermark embedding and extraction procedure is described in detail in following section.

3.3. Watermark embedding procedure

This section discusses the watermark embedding procedure in detail. The watermark is embedded in fifth level detailed coefficients of wavelet transform. In DWT, a one dimensional signal is divided into two parts: high-frequency detailed sub-band and low –frequency approximate sub-band, using wavelet filter. The low frequency part is split again into high and low frequencies. This process is repeated for finite number of times. The original signal is restored back similarly using inverse DWT [22]. The detailed algorithm to generate a watermarked signal, *wat_sig*, by embedding a two dimensional watermark image, *wat_img*, with number of bits equal to *watbits* in the host audio signal, *aud_sig*, is given as follows. Audio signal is divided into frames of equal length using function *enframe*. The function *enframe* basically takes one dimensional audio signal as an input and gives a two dimensional array where each column represents single frame. The number of frames of an audio signal is equal to number of watermarking bits to be embedded in the audio signal. Series of audio samples of a frame are passed through series of filters to calculate its wavelet transform. Consider the series of n samples that are passed through a low pass filter with impulse response, G, resulting in a convolution of the two. The signal is also decomposed simultaneously using a high pass filter, H. The filter outputs are then sub-sampled and give the detail coefficients (from the high-pass filter), Y_{low}, and approximation coefficients (from the low-pass), Y_{high}. This decomposition has halved the resolution since only half of each filter output characterizes the signal. However, each output has half the frequency band of the input so the frequency resolution has been doubled. This decomposition is repeated to further increase the frequency resolution and the approximation coefficients decomposed with high and low pass filters and then down-sampled. In this paper, this decomposition is repeated till fifth level to obtain fifth level detailed wavelet coefficients.

$$wat_img = mat2vec(wat_img)$$

$$watbits = size(wat_img)$$

where, mat2vec is used to convert the two dimensional image in to one-dimensional column vector and the elements of the matrix are taken row-wise.

$$[X_1,\ X_2,\ X_3\ \ldots\ldots X_i] = enframe(aud_sig,\ watbits)$$

$$Y[n] = (X * G)[n] = \sum_{k=-\infty}^{\infty} X_i[k]G[n-k] \tag{7}$$

$$Y_{low}[n] = (X*G)[n] = \sum_{k=-\infty}^{\infty} X_i[k]G[2n-k] \tag{8}$$

$$Y_{high}[n] = (X * G)[n] = \sum_{k=-\infty}^{\infty} X_i[k]H[2n-k] \tag{9}$$

With the sub-sampling operator ↓

$$(Y \downarrow k)[n] = Y[kn] \tag{10}$$

The above summation can be written more concisely as:

$$Y_{low} = (X_i * G) \downarrow 2 \tag{11}$$

$$Y_{high} = (X_i * H) \downarrow 2 \tag{12}$$

This decomposition is repeated further to obtain Y_{high_5}, fifth level detailed coefficients. This vector is then converted into a square matrix. The columns of the matrix are then processed using Gram–Schmidt orthogonalization.

$$Y_{high_5} = \langle a_1|a_2|\cdots|a_n\rangle \tag{13}$$

Define the vector projection:

$$proj_e a = \frac{\langle e, a\rangle}{\langle e, e\rangle}\ e \tag{14}$$

Then,

$$\begin{aligned} u_1 &= a_1, e_1 = \tfrac{u_1}{\|u_1\|} \\ u_2 &= a_2 - (a_2.\ e_1).\ e_1,\ \ e_2 = \tfrac{u_2}{\|u_2\|} \\ u_{k+1} &= a_{k+1} - (a_{k+1}.\ e_1).\ e_1 - \cdots -\ \ (a_{k+1}.\ e_k).\ e_k,\ \ e_{k+1} = \tfrac{u_{k+1}}{\|u_{k+1}\|} \end{aligned} \tag{15}$$

Note that $\|.\|$ is the L_2norm. Rearrange the equations so that a_i's are on left where e_i are unit vectors.

$$\begin{aligned} a_1 &= \ \langle e_1,\ a_1\ \rangle e_1 \\ a_2 &= \ \langle e_1,\ a_2\rangle\ e_1 + \langle e_2,\ a_2\rangle\ e_2 \\ a_3 &= \ \langle e_1,\ a_3\rangle\ e_1 + \langle e_2,\ a_3\rangle\ e_2 + \ \langle e_3,\ a_3\rangle\ e_3 \\ &\vdots \\ a_k &= \ \langle e_1,\ a_k\rangle\ e_1 + \langle e_2,\ a_k\rangle\ e_2 + \langle\ e_3,\ a_k\rangle\ e_3 + \ \cdots + \ \langle e_k,\ a_k\rangle \end{aligned}$$

Where, $\langle e_j,\ a_j\rangle = \|u_j\|$. This can also be written as:

$$\begin{aligned} Y_{high_5} &= \langle a_1|a_2|\cdots|a_n\rangle \\ &= \langle e_1|e_2|\cdots|e_n\rangle \begin{bmatrix} a_1.e_1 & a_2.e_1 & \cdots & a_n.e_1 \\ 0 & a_2.e_2 & \cdots & a_n.e_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n.e_n \end{bmatrix} \end{aligned} \tag{16}$$

$$a_n.e_n'' = \begin{cases} a_n.e_n\ + 2\Upsilon -\ \alpha -\ (\{a_n.e_n\} mod\ \Delta), if\ wat_img_i = 1; \\ a_n.e_n\ + 2\Upsilon -\ \beta -\ (\{a_n.e_n\} mod\ \Delta), if\ wat_img_i = 0 \end{cases} \tag{17}$$

Where, Δis the embedding intensity, α, β and Υ are quantization parameters. The above equation is used to embed the watermarking bit in the audio signal.

$$\begin{aligned} Y_{high_5''} &= \langle a_1|a_2|\cdots|a_n\rangle \\ &= \langle e_1|e_2|\cdots|e_n\rangle * \begin{bmatrix} a_1.e_1 & a_2.e_1 & \cdots & a_n.e_1 \\ 0 & a_2.e_2 & \cdots & a_n.e_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n.e_n'' \end{bmatrix} \end{aligned} \tag{18}$$

$$\begin{aligned} &Y_{(high_5'')} = mat2vec(Y_{(high_5'')}) \\ &X_i'' = \textstyle\sum_{k=-\infty}^{\infty} Y_{low}[k]G[n-2k] + \sum_{k=-\infty}^{\infty} Y_{high_5''}[k]H[n-2k] \\ &wat_sig = \left[X_1''\ X_2''\ X_3''\ \ldots\ldots\ X_i''\right] \end{aligned} \tag{19}$$

The above steps describes the watermark embedding procedure in detail. In this paper, an audio signal is divided into number of frames equal to number of watermarking bits such that one bit can be embedded in one frame to achieve good robustness against signal processing attacks. Number of watermarking bits is evaluated using adaptive payload estimation method proposed in this paper. Audio frames are transformed into wavelet domain and fifth level detailed wavelets coefficient are selected for embedding watermarking bit. Using the high frequency band of fifth level decomposition where human auditory system is not very much sensitive to changes, high embedding capacity has been achieved within perceptual constraints. Thus, this method meets the contradictory design requirements of watermarking algorithm.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

3.4. Watermark extraction procedure

At the receiving end, received watermarked audio is divided into number of frames equal to number of watermarking bits as was done during embedding. The algorithm proposed in this paper can embed different number of watermarking bits in different signals. To meet this requirement of the proposed algorithm, payload is also evaluated at the receiver end when watermarked signal is received. This means that, number of watermarking bits should be known at the receiver end. The detailed procedure is given as follows:

$$[X_1,\ X_2,\ X_3\ \ldots\ldots X_i] = enframe(wat_sig,\ watbits)$$

$$Y[n] = (X * G)[n] = \sum_{k=-\infty}^{\infty} X_i[k]G[n-k] \tag{20}$$

$$Y_{low}[n] = (X*G)[n] = \sum_{k=-\infty}^{\infty} X_i[k]G[2n-k] \tag{21}$$

$$Y_{high}[n] = (X * G)[n] = \sum_{k=-\infty}^{\infty} X_i[k]H[2n-k] \tag{22}$$

$$Y_{high_5} = \langle a_1|a_2|\cdots|a_n\rangle = \langle e_1|e_2|\cdots|e_n\rangle \begin{bmatrix} a_1.e_1 & a_2.e_1 & \cdots & a_n.e_1 \\ 0 & a_2.e_2 & \cdots & a_n.e_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n.e_n \end{bmatrix} \tag{23}$$

$$W_k = \begin{cases} 1, & if\ \{a_n.e_n\} mod\ \Delta > \Upsilon \\ 0, & if\ \{a_n.e_n\} mod\ \Delta \le \Upsilon \end{cases} \tag{24}$$

Where, Δ and Υ are embedding intensity and quantization parameter respectively as determined during embedding. Eq. (24) represents that the proposed extraction procedure is blind as it does not require the host audio at the receiver side. It can be seen from the extraction procedure that original audio signal or watermark is not required during watermark extraction. Therefore, the watermark extraction is blind in nature and thus the presented algorithm.

4. Experimental results

During experimentation, different types of audio samples are used to check the efficiency of the proposed algorithm against various parameters of audio watermarking algorithms. In the experimentation, Δ and $\Upsilon$are considered as 0.08 and 0.06 respectively.

Some of the audio signals which are used in experimental setup are listed in Table 3. The database used for the experimentation of the proposed algorithm consists of approximately 115 audio files. All audio files used for experimentation are of length 10 seconds each. Audio samples including blues, pop, classical, country, folk and some Western and Indian music were used for embedding watermark. Table 3 give specifications like sampling frequency, bitrate, and duration of various audio samples used for experimentation.

Fig. 7 (a) and (b) represents the sample watermarks used in the experimentation of the proposed work. The watermark is resized on the basis of number of bits evaluated to be embedded in the audio signal.

4.1. Evaluation parameters

To analyze the performance of the proposed algorithms following parameters were used for evaluating imperceptibility, robustness and payload:

Table 3
Specifications of audio samples.

S.No.	Audio sample	Sampling frequency	Bit rate	Duration
1	Angrezi Beat	44.1 kHz	16 bps	10 sec
2	Blues1	44.1 kHz	16 bps	10 sec
3	Blues2	44.1 kHz	16 bps	10 sec
4	Classical1	44.1 kHz	16 bps	10 sec
5	Classical2	44.1 kHz	16 bps	10 sec
6	Country1	44.1 kHz	16 bps	10 sec
7	Country2	44.1 kHz	16 bps	10 sec
8	Crazy4u	44.1 kHz	16 bps	10 sec
9	Folks1	44.1 kHz	16 bps	10 sec
10	Folks2	44.1 kHz	16 bps	10 sec
11	Gangsta Zone	44.1 kHz	16 bps	10 sec
12	Give me sunshine	44.1 kHz	16 bps	10 sec
13	I am so lonely	44.1 kHz	16 bps	10 sec
14	I miss her	44.1 kHz	16 bps	10 sec
15	I run to you	44.1 kHz	16 bps	10 sec
16	Kabhi Aditi	44.1 kHz	16 bps	10 sec
17	Love story	44.1 kHz	16 bps	10 sec
18	Mama's song	44.1 kHz	16 bps	10 sec
19	Man I want	44.1 kHz	16 bps	10 sec
20	My world	44.1 kHz	16 bps	10 sec
21	On the floor	44.1 kHz	16 bps	10 sec
22	Pop1	44.1 kHz	16 bps	10 sec
23	Pop2	44.1 kHz	16 bps	10 sec
24	Rhythm divine	44.1 kHz	16 bps	10 sec
25	Smack that	44.1 kHz	16 bps	10 sec
26	You lost me	44.1 kHz	16 bps	10 sec



Fig. 7. (a) Sample watermark image 1, (b) sample watermark image 2.

Table 4
Subjective grades.

Impairment	Quality	Grade
Very annoying	Bad	1
Annoying	Poor	2
Slightly annoying	Fair	3
Perceptible, but not annoying	Good	4
Imperceptible	Excellent	5

1. *Signal to Noise Ratio (SNR)*: The SNR measure is widely used for evaluation of perceptual quality of audio watermarking algorithm, because of its simplicity. SNR determines the difference in decibels between the original and the watermarked audio signal.The value of SNR is evaluated as given in Eq. (2).
2. *Subjective Listening Test*: Like SNR, this test is also used to evaluate the imperceptibility of an audio watermarking algorithm but SNR is an objective quality measure. Subjective grade (SG) is one of the most common methods used for evaluating the quality of watermarked signal. In this test, original signal and the watermarked signals are being played against a set of 10 users out of which 4 are golden ears. These users give a rating to all the watermarked signals as given in Table 4. Then final SG is the mean of scores of all the users.
3. *Normalized Correlation Coefficient (NCC)*: It is a measure used to evaluate the similarity between the embedded and the extracted watermarks. It is given by:

$$NCC = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} I_0(i,j)I_E(i,j)}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} I_0^2(i,j)} \times \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} I_E^2(i,j)}} \tag{25}$$

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

Table 5
Signal to Noise Ratio (SNR).

S.No	Audio sample	Proposed Adaptive Algorithm (payload based on localized property)		Non- Adaptive Algorithm (Uniform Payload)	
		SNR in dB	Payload	SNR in dB	Payload
1	Angrezi Beat	36.92175	702	38.04396	640
2	Blues1	36.38563	702	37.37943	640
3	Blues2	35.13599	384	33.22814	640
4	Classical1	32.17822	384	30.41968	640
5	Classical2	31.16546	318	28.73062	640
6	Country1	37.60783	702	38.52367	640
7	Country2	39.15886	447	34.55973	640
8	Crazy4u	27.40558	384	21.52254	640
9	Folks1	38.34596	768	39.45495	640
10	Folks2	38.24291	768	39.30595	640
11	Gangsta Zone	38.17724	702	39.1404	640
12	Give me sunshine	35.9514	384	34.27723	640
13	I am so lonely	39.01958	447	35.05094	640
14	I miss her	28.67438	318	26.05255	640
15	I run to you	25.78573	384	19.05122	640
16	Kabhi Aditi	35.28608	702	36.0191	640
17	Love story	20.5533	308	14.93024	640
18	Mama's song	21.99282	318	17.43444	640
19	Man I want	31.2511	384	29.3145	640
20	My world	27.35594	384	25.49019	640
21	On the floor	30.66259	384	28.68625	640
22	Pop1	40.68983	831	41.33417	640
23	Pop2	37.42649	702	38.40356	640
24	Rhythm divine	40.4738	831	41.1183	640
25	Smack that	36.9867	702	38.12972	640
26	You lost me	33.372	384	31.40353	640

where, I_o and I_E denote original and extracted binary watermark images. MATLAB inbuilt function corrcoef (x, y) is used to evaluate this parameter.

4. *Bit Error Rate (BER)*: BER is defined as the number of bit errors divided by the total number of transferred bits during interval of time.

$$BER = \frac{\sum_{i=1}^{M}\sum_{j=1}^{M} Io(i,j) \oplus Ie(i,j)}{M \times M} \quad (26)$$

where, I_o and I_E denote original and extracted binary watermark images respectively and $\oplus$ is the exclusive OR operator (XOR).

4.2. Imperceptibility tests

In this paper, proposed algorithm is tested for maintaining perceptual transparency of audio signal under different payload by objective and subjective means. Signal to noise ratio is obtained as an objective measure and mean grades are obtained using subjective listening test.

Table 5 represents the experimental recordings of SNR measure for both adaptive and non-adaptive algorithm. An efficient audio watermarking algorithm should be able to hide watermark data in an audio signal without disturbing the perceptual transparency. An audio with a signal-to-noise ratio greater than or equal to 20 dB can be considered as a good quality signal in which changes are not audible. In this paper, the aim is to achieve optimum payload rate without disturbing the perceptual transparency of the audio signal. The audio signals having less SNR should carry less payload than fixed and the signal having more SNR values are then loaded with more data maintaining perceptual constraints. It can be seen from Table 5 that there are some signals having SNR value less than 20 dB at 640 bps. These audio signals when tested using the proposed adaptive algorithm can produce good quality watermarked signal with SNR value greater than or equal to 20 dB at lower payload rates comparatively. Also there are some signals whose payload have been increased which have very good SNR value at 640 bps. For example, pop1 gives 41.334 SNR at fixed payload which is 640 bps, it is then tested using the proposed adaptive algorithm. It is then found that payload of Pop1 has been increased by 29% without much affecting the perceptual transparency. The payload of pop1 audio signal is 831 bps using adaptive algorithm with good robustness against signal processing attacks.

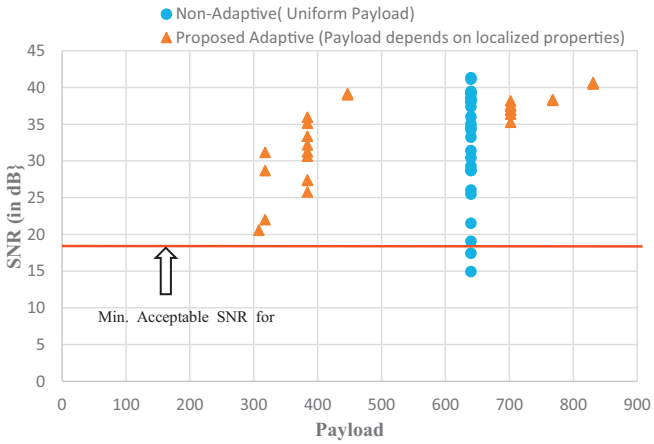


Fig. 8. Plot depicting optimized payload of proposed adaptive algorithm.

Fig. 8 gives the comparison between proposed adaptive and proposed non-adative algorithm diagrammatically in terms of payload and the perceptual transparency. It depicts that the use of proposed adaptive watermarking algorithm gives better SNR values than the fixed payload audio watermarking algorithm. All the audio signals have SNR value greater than 20 dB with optimized payload for proposed adaptive watermarking algorithm. It can be seen from the figure that fixed payload algorithm produces SNR less than the specified standards for some audio signals. So, the payload is optimized using the proposed adaptive audio watermarking algorithm under perceptual constraints.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

Table 6
Subjective listening test.

S.No	Audio sample	Proposed adaptive algorithm (payload based on localized property)		Non adaptive algorithm (uniform payload)	
		SG	Payload	SG	Payload
1	Angrezi Beat	5	702	5	640
2	Blues1	5	702	5	640
3	Blues2	5	384	4	640
4	Classical1	4.5	384	3.5	640
5	Classical2	4.5	318	3	640
6	Country1	5	702	5	640
7	Country2	4.7	447	3.8	640
8	Crazy4u	4.2	384	2.7	640
9	Folks1	5	768	5	640
10	Folks2	4.5	768	5	640
11	Gangsta Zone	5	702	5	640
12	Give me sunshine	4.4	384	3.2	640
13	I am so lonely	5	447	4.5	640
14	I miss her	4.6	318	3.8	640
15	I run to you	4.1	384	1.3	640
16	Kabhi Aditi	5	702	5	640
17	Love story	3.4	308	1.2	640
18	Mama's song	4	318	1.6	640
19	Man I want	4.9	384	4.1	640
20	My world	4.9	384	3.5	640
21	On the floor	5	384	4.6	640
22	Pop1	5	831	5	640
23	Pop2	5	702	5	640
24	Rhythm divine	5	831	5	640
25	Smack that	4.8	702	5	640
26	You lost me	4.5	384	3.7	640

Table 6 gives the results of subjective listening test for adaptive and non-adaptive algorithm. There are some cases, where SNR is above specified standards but SG score is less. For example, the SNR of Classical2 with fixed payload non adaptive algorithm is 28.73 with SG score of 3. This low SG score implies that changes in the audio signal due to the insertion of watermark data are audible to the human ear but scientifically the SNR value is satisfied. The use of proposed adaptive algorithm for embedding watermark after evaluating the payload on the basis of audio signal has improved the SG score with SNR. These audio signals are found to be low energy signals and then their payload is adjusted using the adaptive algorithm to produce good quality watermarked signals. It can be seen from the results that the SG of all the signals when tested using adaptive algorithm is above 4 which means that the changes made to the host audio signal due to insertion of watermark are not audible to human ear. Most of the audio samples have subjective grade of 5 which clearly indicates the exact similarity of original and watermarked audio to human ear.

Hence, the proposed adaptive algorithm is able to keep the perceptual constraint at the acceptable optimized payloads.

4.3. Robustness tests

When the watermarked file is transferred over the internet may undergo some changes due to application of intentional or unintentional signal processing attacks. The effective watermarking algorithm should be able to extract the watermark properly, even if some attacks are applied on it. To check the proposed algorithm against robustness normalized correlation coefficient and bit error rate are evaluated in this paper. The following signal processing attacks are applied on the watermarked audio signals to check the accuracy and robustness of the proposed algorithm:

1. *Mp3 Compression*: The MPEG-1 layer-3 compression is applied on the watermarked signal. The watermarked signal is compressed at different rates i.e. 64 kbps, 128 kbps, 192 kbps and 256 kbps. The compressed signals are then restored back to wave format.
2. *Additive white Gaussian Noise (AWGN)*: White Gaussian Noise is added to the watermarked audio signal so that the attacked signal has a SNR of 30 dB
3. *Resampling*: The watermarked signal is resampled at 22.05 kHz and then resampled back again to 44.1 kHz.
4. *Low pass filtering*: A second order butter worth filter is applied on the watermarked signal with cut-off frequency of 20 kHz.
5. *Re-quantization*: The watermarked audio signal is re-quantized to 8 bits/sample and then back to 16 bits/sample.

Table 7
Extracted watermark without attack for blues1 audio.

No attack	Payload	NCC	BER
Adaptive	702	0.999495	0.000142
Extracted watermark sample 1			
Extracted watermark sample 2			
Non-adaptive	640	0.999686	0.000156
Extracted watermark sample 1			
Extracted watermark sample 2			

Table 7 gives the extracted watermark from Blues1 audio without attack for both adaptive and non-adaptive algorithms. It can be seen from the given Table that, if the payload has been increased for blues1 audio using adaptive algorithm, the identity of the extracted watermark can be recognized properly. The proposed algorithm is tested against robustness by using adaptive algorithm and by fixing payload at 640 bps.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

Table 8
Experimental results for robustness.

Audio sample	Attack type	Proposed adaptive algorithm (payload based on localized property)		Non- adaptive algorithm (uniform payload)	
		NCC	BER	NCC	BER
Angrezi Beat	No attack	0.999495	0.000142	0.999686	0.000156
	Mp3 Comp (64 kbps)	0.999495	0.000142	0.985635	0.004427
	Mp3 Comp (128 kbps)	0.999495	0.000142	0.985635	0.004427
	Mp3 Comp (192 kbps)	0.999495	0.000142	0.985635	0.004427
	Mp3 Comp (256 kbps)	0.999495	0.000142	0.999686	0.000156
	AWGN	0.890357	0.054531	0.773027	0.07792
	Re-quantization	0.925837	0.036875	0.853757	0.047009
	Resampling	0.999686	0.000156	0.99949	0.000142
	LPF	0.793419	0.102656	0.833016	0.054701
Blues1	No attack	0.999058	0.000469	1	0
	Mp3 Comp (64 kbps)	0.999058	0.000469	1	0
	Mp3 Comp (128 kbps)	0.999058	0.000469	1	0
	Mp3 Comp (192 kbps)	0.999058	0.000469	1	0
	Mp3 Comp (256 kbps)	0.999058	0.000469	1	0
	AWGN	0.85103	0.074063	0.982322	0.005469
	Re-quantization	0.885922	0.056719	0.992984	0.001994
	Resampling	0.999058	0.000469	0.999058	0.000469
	LPF	0.797821	0.100469	0.982322	0.005469
Blues2	No attack	1	0	0.964531	0.011198
	Mp3 Comp (64 kbps)	1	0	0.9978	0.001094
	Mp3 Comp (128 kbps)	1	0	0.964531	0.011198
	Mp3 Comp (192 kbps)	1	0	0.952171	0.01288
	Mp3 Comp (256 kbps)	1	0	0.964531	0.011198
	AWGN	0.667261	0.165469	0.41139	0.264323
	Resampling	1	0	0.964531	0.011198
	LPF	1	0	0.987335	0.003906
	Requantization	0.716672	0.140938	0.49916	0.215885
Classical1	No attack	1	0	0.982322	0.005469
	Mp3 Comp (64 kbps)	1	0	0.924693	0.0375
	Mp3 Comp (128 kbps)	1	0	0.924693	0.0375
	Mp3 Comp (192 kbps)	1	0	0.982322	0.005469
	Mp3 Comp (256 kbps)	1	0	0.982322	0.005469
	AWGN	0.763597	0.1175	0.527181	0.195573
	Resampling	1	0	0.982322	0.005469
	LPF	0.659978	0.169063	0.695191	0.151406
	Requantization	0.789751	0.104531	0.594887	0.163542
Classical2	No attack	1	0	1	0
	Mp3 Comp (64 kbps)	1	0	1	0
	Mp3 Comp (128 kbps)	1	0	1	0
	Mp3 Comp (192 kbps)	1	0	1	0
	Mp3 Comp (256 kbps)	1	0	1	0
	AWGN	0.793419	0.102656	0.621201	0.149686
	Resampling	1	0	1	0
	LPF	0.754468	0.122031	0.61895	0.151042
	Requantization	0.833096	0.082969	0.694149	0.11195

(*continued on next page*)

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

Table 8 (*continued*)

Audio sample	Attack type	Proposed adaptive algorithm (payload based on localized property)		Non- adaptive algorithm (uniform payload)	
		NCC	BER	NCC	BER
Country1	No attack	0.992984	0.001994	0.9978	0.001094
	Mp3 Comp (64 kbps)	0.934092	0.032813	0.999686	0.000156
	Mp3 Comp (128 kbps)	0.934092	0.032813	0.999686	0.000156
	Mp3 Comp (192 kbps)	0.9978	0.001094	0.999686	0.000156
	Mp3 Comp (256 kbps)	0.9978	0.001094	0.999686	0.000156
	AWGN	0.857407	0.070938	0.773607	0.077778
	Resampling	0.9978	0.001094	0.992984	0.001994
	LPF	0.617511	0.190313	0.766651	0.115938
	Requantization	0.905405	0.047031	0.833016	0.054701
Country2	No attack	0.999686	0.000156	0.617511	0.190313
	Mp3 Comp (64 kbps)	0.999686	0.000156	0.617511	0.190313
	Mp3 Comp (128 kbps)	0.999686	0.000156	0.617511	0.190313
	Mp3 Comp (192 kbps)	0.999686	0.000156	0.617511	0.190313
	Mp3 Comp (256 kbps)	0.999686	0.000156	0.617511	0.190313
	AWGN	0.766651	0.115938	0.003919	0.525
	Resampling	0.999686	0.000156	0.002677	0.529514
	LPF	0.634013	0.182188	-0.00476	0.522396
	Requantization	0.821428	0.08875	0.000742	0.529688
Crazy4u	No attack	0.999686	0.000156	0.973268	0.008333
	Mp3 Comp (64 kbps)	0.875198	0.062031	0.735556	0.098698
	Mp3 Comp (128 kbps)	0.999686	0.000156	0.964531	0.011198
	Mp3 Comp (192 kbps)	0.999686	0.000156	0.973268	0.008333
	Mp3 Comp (256 kbps)	0.999686	0.000156	0.973268	0.008333
	AWGN	0.617511	0.190313	0.364323	0.28724
	Resampling	0.999686	0.000156	0.973268	0.008333
	LPF	0.754468	0.122031	0.59385	0.163802
	Requantization	0.634013	0.182188	0.431154	0.250781
Folks1	No attack	0.998743	0.000625	0.978215	0.006771
	Mp3 Comp (64 kbps)	0.983974	0.007969	0.978215	0.006771
	Mp3 Comp (128 kbps)	0.998743	0.000625	0.978215	0.006771
	Mp3 Comp (192 kbps)	0.998743	0.000625	0.978215	0.006771
	Mp3 Comp (256 kbps)	0.998743	0.000625	0.978215	0.006771
	AWGN	0.878868	0.060313	0.982236	0.005469
	Resampling	0.998743	0.000625	0.999686	0.000156
	LPF	0.709697	0.108073	0.86482	0.067188
	Requantization	0.924355	0.037656	0.978215	0.006771
Folks2	No attack	0.998115	0.000938	0.998115	0.000938
	Mp3 Comp (64 kbps)	0.934092	0.032813	0.97316	0.000512
	Mp3 Comp (128 kbps)	0.985617	0.007188	0.998115	0.000938
	Mp3 Comp (192 kbps)	0.985617	0.007188	0.999057	0.000469
	Mp3 Comp (256 kbps)	0.985617	0.007188	0.999057	0.000469
	AWGN	0.881365	0.059063	0.986467	0.004167
	Resampling	0.998115	0.000938	0.998115	0.000938
	LPF	0.735556	0.098698	0.89234	0.05375
	Requantization	0.921906	0.038906	0.998487	0.000427

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

Table 9
Comparison of proposed algorithm with some existing algorithms.

Algorithm	Payload (bps)	SNR (dB)	BER (%)				
			Noise addition	Resampling	Mp3 compression	LPF	Re-quantization
Hemis et al. [23]	102.4	50.445	0	0 (22.05)	0.36 (128 kbps)	0.06 (11.05	0 (8 bits/ sample)
Ali Al-Haj [24]	258	38.17	0 (20 dB)	0.0128 (22.05)	0.0727 (64 kbps)	0.1893 (8 kHz)	0 (8 bits/ sample)
Yang et al. [25]	NR	NR	0.316 (100)	0.1333 (22.05)	0.1167 (64 kbps)	0 (8 kHz)	NR
Noriega et al. [26]	229	> 40	NR	NR	0.01.323	NR	NR
Bhat et al. [27]	196	NR	0 (20 dB)	0 (22.05)	0.01 (64kbps)	0 (11.025 kHz)	0 (8 bits/ sample)
Proposed (Non-adaptive)	**640**	**37.379**	**0.264323 (30 dB)**	**0.011198 (22.05)**	**0.001094 (64, 128 kbps)**	**0.182188 (20 kHz)**	**0.215885 (8 bits/ sample)**
Proposed (Adaptive)	**702**	**36.385**	**0.165469 (30 dB)**	**0 (22.05)**	**0 (64, 128 kbps)**	**0 (20 kHz)**	**0.140938 (8 bits/ sample)**

NR: Not Reported

Table 8 clearly indicates the robustness of proposed algorithm under different signal processing intentional or unintentional attacks for both adaptive and non-adaptive algorithms. Table 8 gives the results of some sample audio signals tested for robustness against additive white Gaussian noise, resampling, low-pass filtering and mp3 compression at 64 kbps, 128 kbps, 192 kbps and 256 kbps. In the above Table results for only few audio signal has been presented due to space limitation. It can be seen from the experimental results that the proposed adaptive algorithm exhibits very good robustness against different audio attacks even the payload has been has been increased for a particular audio sample.

Table 9 gives the comparison of the proposed algorithm with some of the existing audio watermarking algorithms in literature. The proposed algorithm have shown good robustness results in terms of BER values equal to or closer to zero at such a high payload. Also proposed algorithm has shown good performance in terms of imperceptibility with comparison to techniques existing in the literature. It can be seen from the table that the proposed algorithm is able to achieve payload of 702 bps with good robustness against signal processing attacks and good perceptual transparency.

5. Conclusion

In this paper, an adaptive audio watermarking algorithm is proposed, which is designed to evaluate payload of audio signal on the basis of localized audio features. The proposed adaptive algorithm aims to resolve the problem of over-loading and under-loading the audio signals with watermark data making the payload optimized for each individual audio host signal. PCA is used in this paper for finding the discriminatory features which has relevance with the payload of the audio signal. A mathematical model is designed using the selected audio features like energy, zero cross mean and short time energy to efficiently evaluate the degree of embedding under perceptual transparency. Experimental results indicate that the proposed adaptive algorithm provide good imperceptibility with good robustness against signal processing attacks at adjustable payload for different types of audio signals. Comparative analysis reveals that this proposed adaptive algorithm has better performance than the fixed payload non-adaptive algorithm.

References

[1] Bender W, Gruhl D, Morimoto N. Techniques for data hiding. IBM Syst J 1996;35(¾):131–336.
[2] Hua Guang, Goh J, Thing VLL. Time-spread echo-based audio watermarking with optimized imperceptibility and robustness. IEEE/ACM Trans Audio Speech Lang Process 2015;23(2):227–39.
[3] Dutta MalayKishore, Gupta Phalguni, Pathak VinayK. A perceptible watermarking algorithm for audio signals". In: Multimedia tools & applications. Springer; 2014. p. 691–713.
[4] Erfani Y, Siahpoush S. Robust audio watermarking using improved TS echo hiding. In: Digital signal processing, 19. Elsevier; 2009. p. 506–16.
[5] Kaur Arashdeep, Dutta MK, Soni KM, Taneja Nidhi. A blind audio watermarking algorithm robust against synchronization attacks. In: International conference on signal processing, computing and control (ISPCC); 2013. p. 1–6.
[6] Sourabh Dimri, Singh Sudhir, Kaur Arashdeep, Dutta MK. A Robust watermarking algorithm based on multi resolution decomposition of audio signal. In: 2012 Third international conference on computer and communication technology; 2012. p. 299–302.
[7] Li Xueyao, Zhang Min, Sun Shiliang. Adaptive audio watermarking algorithm based on SNR in wavelet domain. In: International conference on natural language processing and knowledge engineering; 2003. p. 287–92.
[8] Pooyan M, Delforouzi A. Adaptive and robust audio watermarking in wavelet domain. In: Third International conference on intelligent information hiding and multimedia signal processing; 2007. p. 287–90.
[9] Wang Xun, Peng Hong, Hu Deyu, Wang Jun. A robust watermarking scheme based on audio features in multiwavelet domain. In: Pacific-Asia workshop on computational intelligence and industrial application; 2008. p. 408–12.
[10] Li Dahui. A new algorithm of self-adaptability for digital audio watermark with energy quantization. In: Pacific-Asia conference on circuits, communications and systems; 2009. p. 559–62.
[11] Dutta MalayKishore, Pathak VinayK, Gupta Phalguni. An adaptive robust watermarking algorithm for audio signals using SVD. In: Transactions on computational science. Springer; 2010. p. 131–53.
[12] Youssef SherinM. HFSA-AW: a hybrid fuzzy self-adaptive audio watermarking. In: International conference on communications, signal processing and their applications; 2013. p. 1–6.
[13] Peng Hong, Wang Jun, Wang Weixing. An adaptive audio watermarking method based on local audio feature and support vector regression. In: International conference on software engineering, artificial intelligences, networking and parallel/distributed computing; 2009. p. 381–4.
[14] Hu HwaiTsu, Hsu LingYuan, Chou HsienHsin. Perceptual based DWPT-DCT framework for selective blind audio watermarking. In: Signal processing, 105. Elsevier; 2014. p. 316–27.
[15] Fallahpour M, Megías D. High capacity method for real-time audio data hiding using the FFT transform. In: Advances in information security and its application. Springer-Verlag; 2009. p. 91–7.
[16] James Leonard Fiocca, "Method and apparatus of reducing processing steps in an audio compression system using psychoacoustic parameters", U.S. Patent US5732391 A, March 24, 1998.
[17] Yang X, Tan B, Ding J, Zhang J, Gong J. Comparative study on voice activity detection algorithm. In: International conference on electrical and control engineering; 2010. p. 599–602.
[18] Enqing D, Guizhong L, Yatong Z, Yu c. Voice activity detection based on short--time energy and noise spectrum adaptation. In: 6th International conference on signal processing; 2002. p. 464–7.
[19] Jalil Madiha, Butt FaranAwais, Malik Ahmed. Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals. In: International conference on technological advances in electrical, electronics and computer engineering; 2013. p. 208–12.
[20] Coates RM. The law. In: Newman James R, editor. The world of mathematics. New York: Simon and Schuster; 1956. p. 305–24.
[21] Ryan MichaelJohn, Frater Michael. Communications and information systems. Argos Press; 2002. p. 5–10.
[22] Huang Huang-Nan, Chen Shuo-Tsung, Lin Muh-Shi, Kung Woon-Man, Hsu Chih-Yu. Optimization-based embedding for wavelet-domain audio watermarking. J Signal Process Syst 2013;80(2):197–208 Springer.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

[23] Hemis Mustapha, Boudraa Bachir, Merazi-Meksen Thouraya. Optimized audio watermarking scheme with swarm intelligence. In: First international conference on new technologies of information and communication (NTIC); 2015. p. 1–6.
[24] Al-Haj Ali. An imperceptible and robust audio watermarking algorithm. EURASIP J Audio Speech Music Process 2014(1):1–12.
[25] Yang H, Bao D, XWang PNiu. A robust content based audio watermarking using UDWT and invariant histogram. Multimedia Tools Appl J 2010;57(3):453–76 Springer.
[26] Noriega RM, Nakano M, Kurkoski B, Yamaguchi K. High payload audio watermarking: toward channel characterization ofMP3 compression. J Inf Hiding Multimedia Signal Process 2011;2(2):91–107.
[27] Vivekananda Bhat K, Sengupta Indranil, Das Abhijit. An audio watermarking scheme using singular value decomposition and dither-modulation quantization. Multimedia Tools Appl J 2011;52(2-3):369–83.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003



Kaur A. received B.Tech & M.Tech in CSE from PTU, Jalandhar, India in 2006 and 2008 respectively. She is currently pursuing Ph.D. in CSE at Amity University, Noida, India.
Since 2008, she is with CSE department, Amity University. Her research interest includes audio watermarking, software reliability and image processing. She has published various peer reviewed papers in International Conferences and journals.



M.K. Dutta received M.Tech in ECE from Central University, Tezpur India. He did his PhD in Multimedia Data security and Signal Processing from UPTU.
Since 2010, he is associated with the ECE department, AUUP, India. His interests include data security, watermarking and steganography. He has published many peer-reviewed papers in International journals and conferences. He has filed 3 patents and is principal investigator of various funded projects.



Dr. K.M Soni obtained his B.E. (Electrical Engineering) and M.E. (Control and Instrumentation) degrees from Motilal Nehru Regional Engineering College (Presently known as Motilal Nehru National institute of Technology) Allahabad (U.P) and his Ph. D. from Jamia Millia Islamia (a Central University) New Delhi. He is currently associated with Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida (U.P) and has been involved with teaching the students in the area of Circuits / Networks and Systems, Signal and Systems, Digital Signal Processing, Control Systems, Modern/Digital Control Systems, Process Control, Power Electronics, Electrical Science and Machines etc. at undergraduate and postgraduate levels. He has also authored the books on Signals and Systems, Basic System Analysis, Network Analysis and Synthesis, and Advanced Control Systems. He has also supervised many B. Tech and M. Tech projects and presented/published technical papers in International Conferences and Journals. He is Life Member of Indian Society for Technical Education (ISTE), Institution for Electronics and Telecommunication Engineers (IETE), International Journal of Engineering Research and Industrial Applications (IJERIA), Member of IEEE, IET(U.K), CSI and other similar Organization.



Nidhi Taneja is associated with the Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technological University for Women, DELHI. She did her Ph.D. from IIT, Roorkee in the field of Security solutions for still Visual Data. She has published many peer-reviewed research papers in International journals and conferences.

Please cite this article as: A. Kaur et al., Localized & self adaptive audio watermarking algorithm in the wavelet domain, Journal of Information Security and Applications (2017), http://dx.doi.org/10.1016/j.jisa.2016.12.003

This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

Volume 77 • Issue 1 • July 2016

Journal of Achievements in Materials and Manufacturing Engineering

International Scientific Journal published monthly by the World Academy of Materials and Manufacturing Engineering

Mechanical properties and microstructure studies in Friction Stir Welding (FSW) joints of dissimilar alloy – a review

H. Mehdi [a,*], R.S. Mishra [b]

[a] PhD Scholar, Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

[b] Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

* Corresponding e-mail address: husainmehdi4u@gmail.com

ABSTRACT

Purpose: Friction stir welding (FSW) is a relatively new solid state joining process that uses a non-consumable tool to join two different material without melting the workpiece material. Friction stir welding (FSW) was developed for microstructural modification of metallic material. This review article provides an overview of effect of FSW/FSP mechanism responsible for the formation of weld, microstructure refinement, wear of FSW tool and mechanical properties. This review conclude with recommendations for future research direction.

Design/methodology/approach: Heat is generated by friction between the rotating tool and the workpiece material. This joining process is energy efficient, environment friendly and versatile.

Keywords: Friction Stir Welding; Microstructure; Hardness; Tensile strength; Tool wear

Reference to this paper should be given in the following way:

H. Mehdi, R.S. Mishra, Mechanical properties and microstructure studies in Friction Stir Welding (FSW) joints of dissimilar alloy – a review, Journal of Achievements in Materials and Manufacturing Engineering 77/1 (2016) 31-40.

PROPERTIES

1. Introduction

Friction Stir Processing (FSP)/FSW is a method of changing the properties of a metal through intense localized plastic deformation. This deformation is produced by forcibly inserting a non-consumable tool into the workpiece and revolving the tool in a stirring motion as it is pushed laterally through the workpiece. The antecedent of this technique, friction stir welding is used to join multiple piece of metal without creating the heat affected zone typical of fusion welding (Fig. 1). Efficient joints in terms of strength of aluminium matrix composite materials cannot be achieved by fusion based welding method due to the reaction between reinforcements and matrices leading to the formation of brittle secondary phase in the weld pool or decomposition of reinforcements on molten metal [1,2].

© Copyright by International OCSCO World Press. All rights reserved. 2016

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

As a versatile material, aluminium matrix composites may be selected as an alternative to high strength aluminium alloys in aero engines and aerospace structures like fins, wing and fuselage. In 2001 NASA used composite aluminium AL-Li 2195 rather than aluminium alloy Al2219 for the external fuel tank of space shuttles leading to a reduction of weight by 3400 kg. The saving in weight increases the cargo capacity of space shuttles and enables it to transport more than one components in a single flight to the international space station [3]. Titanium alloy are used extensively in the aerospace industry due to their excellent structure efficiency and good high temperature strength. Welding is an effective way to produce a structure with complex geometry and multiple components. Titanium alloys are readily fusion weld able. However, some problems associates with fusion welding of titanium alloys include porosity, distortion and formation of coarse cast grain structure [4,5].

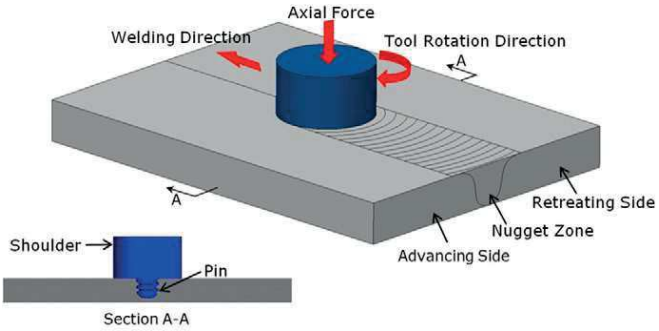


Fig. 1. Friction Stir Welding

2. Literature review

A large number of tool deformations occurred during the first 3 inch length welding trial, the tool configuration changed slightly due to reduced stress on the deformed pin part and tool size and weight decreased continuously. The stress induced cracks were responsible for the majority of tool weight loss.

The defect were found in a region with bimodal microstructure of α+ transformed β phase, which indicate local processing temperature below β- transus due to low heat input [6]. The heat affected zones of a friction stir weld of aluminum alloy 7050-T651 were investigated and compare with the unaffected base metal. Composition of 7050-T651 are 5.7-6.7 Zn, 1.9-2.6 Mg, 2.0-2.6 Cu, 0.08-0.115 Zr. The rotation speed of pin was 350 rpm and travel speed was 15mm/min. Compared to parent material microstructure, the strengthening precipitates have coarsened severely and the precipitate free zone along the grain boundaries has increased by factor of five during friction stir welding, The original base metal grains structure is completely eliminated and replaced by a very fine equiaxed grain structure in the dynamic re-crystallized zone (DXZ). The DXZ consisted of re-crystallized, fine equiaxed grains on the order of 1-4 μm in diameter. Most of the DXZ grains contained a high dislocation density with various degree of recovery from grain to grain [7]. Friction stir processing has been successfully used formation of nano grains and increase the mechanical properties i.e. surface hardness, wear resistance, tensile and fatigue strength.

It was observed that when there is increment on travelling speed, hardness value will also be increased. However increased rotation speeds resulted in lower hardness value at the same travelling speed. Processing parameter including tilt angle and target depth are crucial produce sound and depth free processed region. Friction stir processing result insignificant temperature rise with in and around the weld. A temperature rise of 400-500°C has been recorded within the stir zone for aluminium alloy. The temperature rise result in significant micro structural evaluation i.e. fine re-crystallized grains of 0.1-18 mm, texture, precipitate dissolution and coarsened and residual stress with a much lower magnitude [8].

2.1. Mechanical properties of FSW joints

The investigation of mechanical properties including tensile strength, hardness, and fatigue strength is important particularly for critical component. It is also possible to optimize the welding parameter based on the evaluation of these properties which reflect the joint efficiency.

Tensile strength

Friction stir welding technology requires a thorough understanding of the process and consequent mechanical properties of the weld in order to be used in the production of component for aerospace application. For this reason, detailed research of friction stir welding is required. Friction stir welding can be used to join a different member of material, the primary research and industrial interest has been join aluminium alloy. Defect free welds with good mechanical properties have been made in a wide variety of aluminium alloys, thickness from 1 mm to more than 35 mm will not be welded by friction stir welding. In addition, friction stir weld can be accomplished in any position. [9-15]. The ultimate tensile strength and hardness

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

of bimetallic weld joint increases by increasing the pre-stress, and ductility was decreases when thermal loading increases. For preventing brittle failure behaviour of carbon steel the value of pre-stress and thermal stress should be low as possible [16,17]. The stress strain behaviour of friction stir processing Al-4Mg-1Zr as shown in Figure 2. The optimum strain rate for maximum elongation at 525°C was $1x10^{-1}$ s^{-1}. This show that high strain super plasticity can be achieved in the Al-4Mg-1Zr alloy by FSP.

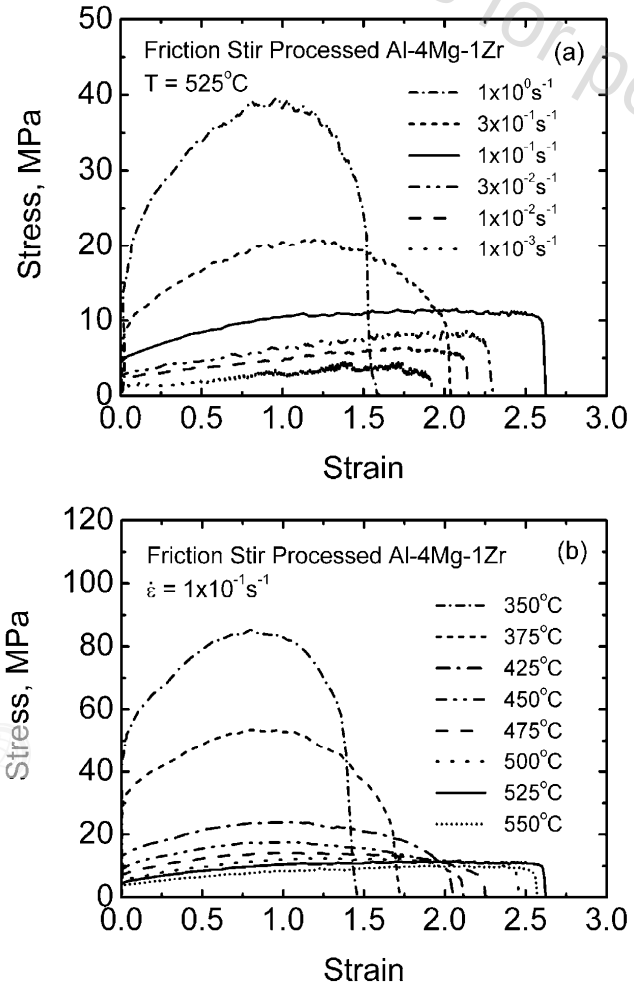


Fig. 2. Stress strain behaviour of friction stir processing Al-4Mg-1Zr, (a) initial strain rate at 525°C, (b) temperature at an initial strain rate $1x10^{-1}$ s^{-1} [18]

Superplastic deformation behaviour of friction stir processing Al-4Mg-1Zr alloy was investigated in strain rate range of $1x10^{-3}$ to 1 s^{-1} and temperature range of 350°-550° and compared with that of as-rolled one. The maximum elongation of 1280% was obtained at 525°C. FSP resulted in significant decrease in the flow stress in Al-4Mg-1Zr alloy. The tested specimen of friction stir processing Al-4Mg-1Zr alloy deformed to failure at 525°C for different strain rates and at $1x10^{-1}$ s^{-1} for different temperature. The specimen show neck free elongation that is characteristic of superplastic flow. The strain rate sensitivity of FSP Al-4Mg-1Zr increases with increasing initial strain rate at investigated strain rates > $1x10^{-3}$, whereas the as-rolled alloy exhibited a nearly constant strain rate sensitivity throughout the investigated strain rate range at 450° C FSP Al-4Mg-1Zr exhibited an increasing m value of 0.12-0.55 in the initial strain rate ranges of 10^{-3} to 10^{-1} s^{-1}. Similarly, the strain rate range sensitivity of FSP Al-4Mg-1Zr at 525°C increases from 0.15 at low strain rates to 0.53 at high strain rates as shown in Figure 2 (a) [18].

The tensile properties of the as-extruded material and the weld with different FSW processing parameters are shown in Figure. 3. The ductility of weld 1 was improved to 11.6% with a slight decrease in strength compared the as-extruded material. A further enhancement in ductility to 14.7% was observed in weld 2, with an accompanying loss of yield strength from 633 MPa to 568 MPa. This decrease in strength is likely to be a result of microstructure coarsening.

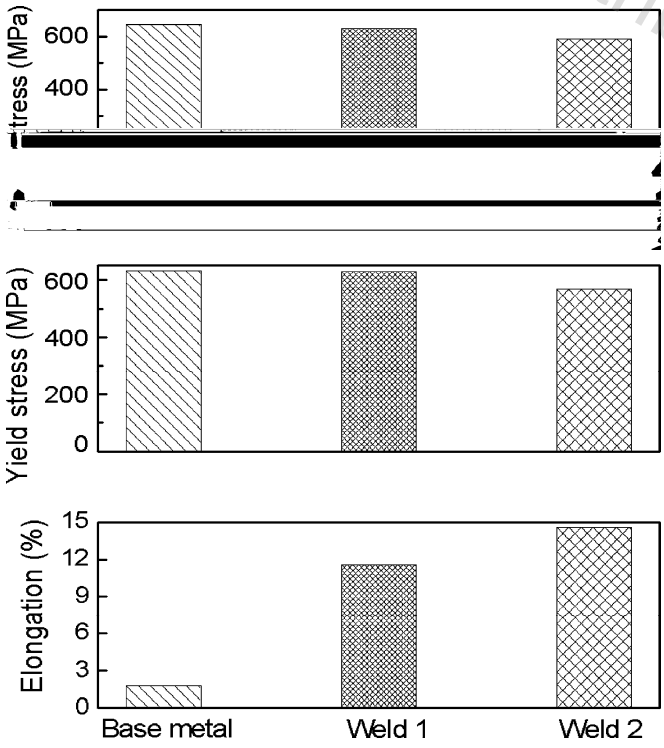


Fig. 3. Tensile properties of as-extruded material weld 1 and weld 2 at room temperature [18]

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

2.2. Microstructure of friction stir welding joint

The basic concept of friction stir welding is remarkably simple. A rotating tool with pin and shoulder is inserted in a single piece of material for microstructure modification and transversed along the desired line to cover the region of interest. Friction between the tool and workpieces resulted in localized heating that softens and plasticizes the workpiece. A volume of processed material is produced by movement of material from the front of the pin to the back of the pin. During this process, this results in significant grain refinement [19,20].

The presence of a fine grained microstructure is a critical criterion for super plasticity friction stir processing has been shown to have potential as a technique for the production of superplastic aluminium alloy. However, the extent of super plasticity in friction stir processing parts is limited at elevated temperatures due to the evolution of a very coarse-grained microstructure [21-24].

The FSP is very effective technique to fabricate surface metal matrix composite with well distributed particles and very good bonding with metal substrate. The advantages of the FSP are evident compared with laser processing, high energy electron beam irradiation and casting sinter [25].

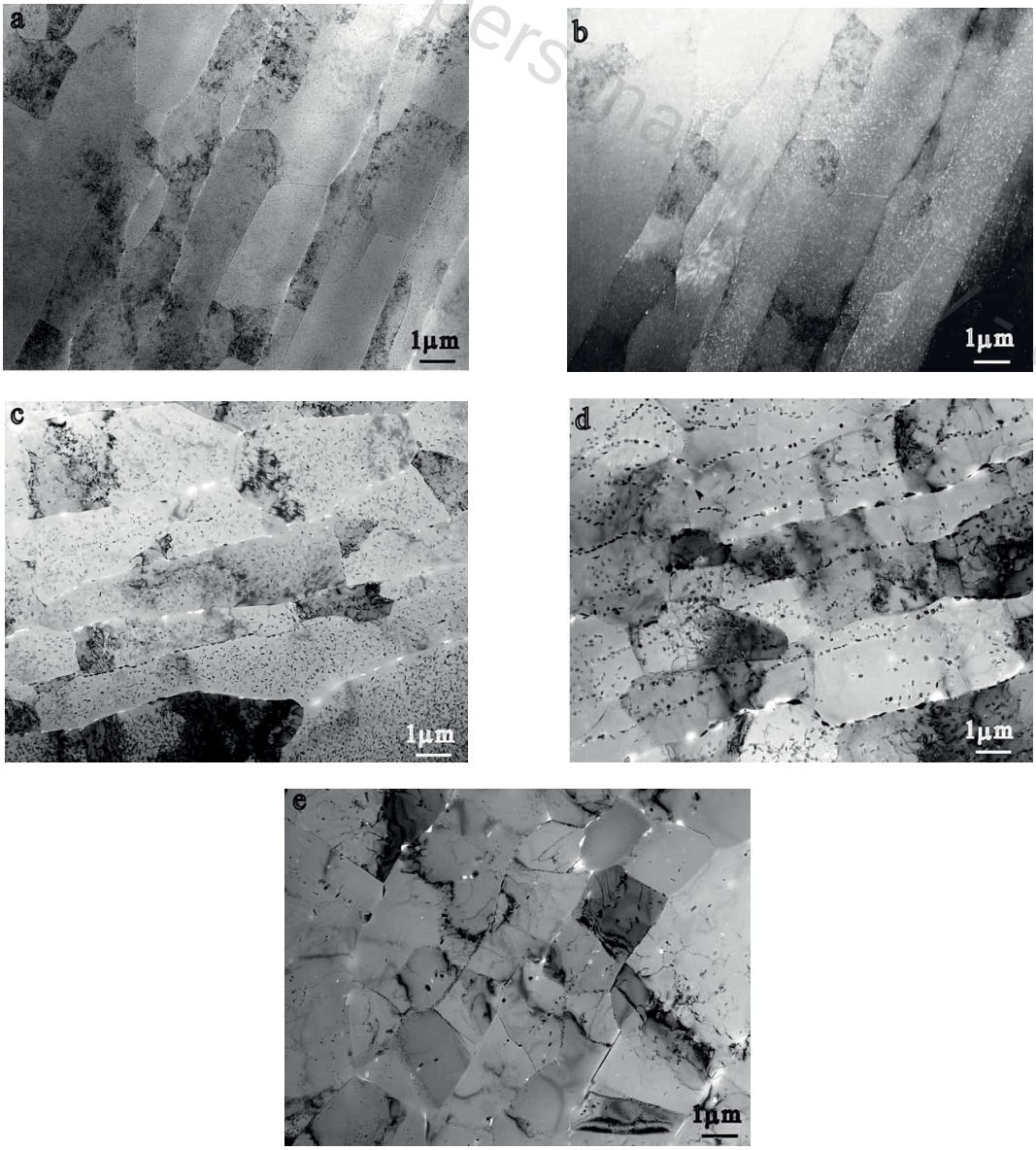


Fig. 4. Grain structure in different weld regions: (a) base metal, (b) dark field image of region (a), (c) HAZ, (d) TMAZ I, (e) TMAZ [26]

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

The grain structure in the base metal and weldment are shown in low magnification transmission electron microscopy (TEM) photo graph as shown in Figure 4. The microstructure in the parent alloy shows partially recrystallized pancake shaped grains with un-recrystallized region containing sub grains of about 1-5 µm (Fig. 4a). The small difference in diffraction contrast between neighbouring grains in the dark field image shows the dominance of low angle boundaries (Fig. 4b). The grains structure in the HAZ region, which has not been disturbed mechanically by Friction stir welding (FSW), is similar to that of the base metal, (Fig.4c). Grains in base metal and HAZ contain a relatively low dislocation density. The thermos-mechanically affected zone (TMAZ), located between the parent metal and the dynamically recrystallized zone (DXZ), is characterized by a highly deformed structure. With coursing of sub grains, the elongated base metal grains have been pre-served in the TMAZ (Fig. 4d) [26].

The influence of process parameter and FSP run configuration on the stability of nugget microstructure at elevated temperature has been evaluated. All single pass runs showed some extent of abnormal gain growth (AGG), whereas multi pass runs were more resistance to AGG. Cast Al-Alloy of F357 we used for this study. This alloy belongs to the hypoeutectic family of Al-Si system. The occurrence of abnormal grain growth increase when the tool rotation rate is reduced from 2236 rpm to 1500 rpm. The most notable feature of this investigation is the observation of change in the microstructural response of the nugget towards AGG as a function of number of passes inside the nugget. The multiple pass does not resulted in Si particles refined beyond a certain limit. The multi pass run of second configuration indicate that the extent of AGG can be reduced if the material is FSPed multiple times.

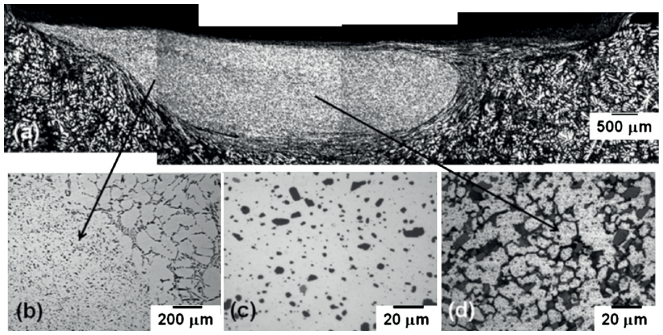


Fig. 5. (a) Cross-section optical macrograph of a single pass run, (b) interface between cast and FSPed region showing extent of particle refinement, (c) higher magnification image shows size and shape of Si particles inside the FSPed nugget and (d) grain structure inside the nugget [27]

Fig. 5a shows an optical macrograph of a single pass run cross-section in which a typical basin shaped nugget can be clearly identified. Eutectic Si particles inside the nugget went through a complex material flow and as a result were highly refined. The extent of refinement is apparent in Fig. 5b, which shows the interface of the cast and FSPed region. The size and shape of refines Si particles get clearer in the higher magnification image Fig. 5c. The typical grain structure of the FSPed nugget is shown in Fig. 5d. The grains were mostly equiaxed, although in a few instance some grains aligned parallel to either the advancing or retreating side [27].

2.3. Tool performance evaluation

Friction stir welding tool wear is a critical issue for aluminium matrix composites, which occurs as a result of friction, rotation and movement of friction stir welding tool along the base material. Plastic deformation abrasion diffusion and reaction between the environment and the tool material are the major wear mechanism that happen in FSW tools [28]. Friction stir welding of soft metal such as aluminium and magnesium did not exhibits significant wear of the tool. However tool life issue becomes more significant when hard metals of high melting temperature or metal matrix composite (MMC) are welded by FSW. His phenomenon is characterized by the deformation and reduction in the pin diameter [29].

Figure 6 shows optical microscopy image of cross section of CY16, W-La and WC411. A crack initiated at shoulder corner and then penetrated through the pin in CY16 tool as shown in fig.6a, crack bifurcation took place during propagation and resulted in rapid fracture and chipping. Further investigation revealed that shoulder edge and shoulder-pin corner were preferred crack initiation sites, which might be due to stress concentration at these regions. Adhesion occurred between Ti-6Al-4V and CY16 as shown in Fig.6b. Tool material cracking under the adhesion layer suggested stress concentration caused by pulling force applied by adhered material. Plastic deformation of W-La tool was revealed by the deformed grain structure in Fig. 6c. The formation of stress induced crack at pin tip could reduce the resistance to wear. No adhesion of Ti-6Al-4V material was observed on W-La tool surface. In the WC411 tool, voids are nucleated at shoulder surface and shoulder corners as shown in Fig. 6e and void coalescence happens due to localized strain. Adhesion layer of Ti-6Al-4V material was also observed on the surface of WC411 tool as shown in Figure 6 [30].

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

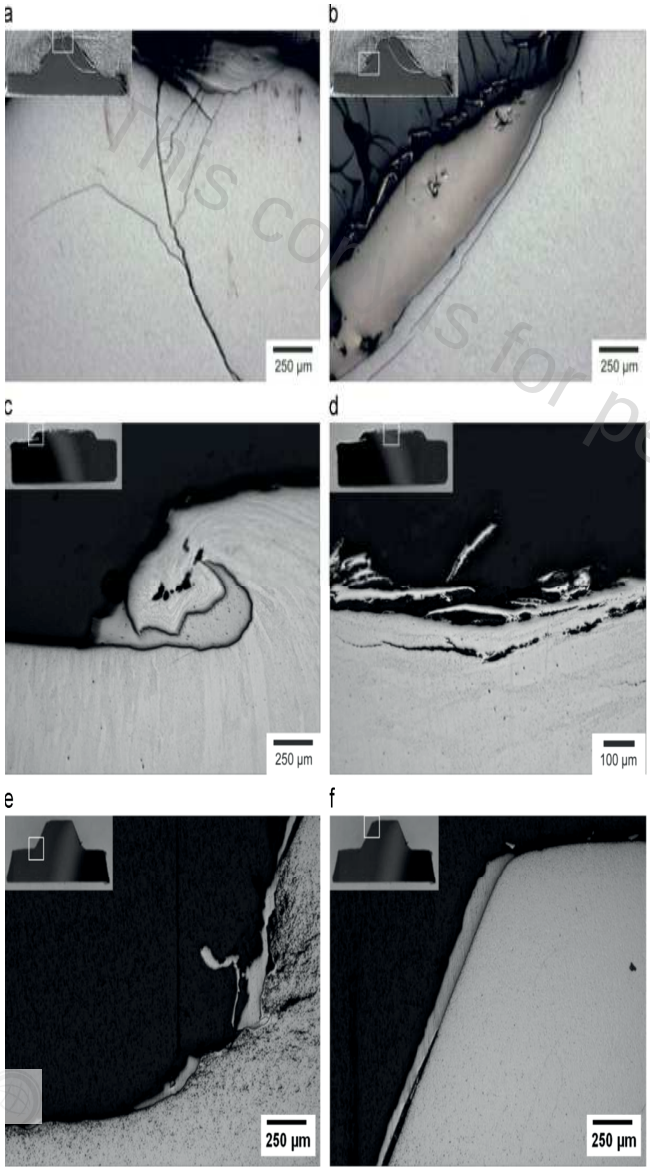


Fig. 6. Optical microscopy images of cross sections of (a, b) CY16, (c, d) W-La, and (e, f) WC411 tools. The inset at left-top corner of each figure. Shows a low magnification image [30]

There was a linear relationship between the wear rate and tool rotation speed. However, as a result of filling threads by base materials the initial wear of the tool can be delayed by the reduction of rotation speed and increase in transverse speed and after a period of time the wear rate becomes constant as shown in Fig. 7. They concluded that the self-optimized tool shape was evolved as a result of complex solid state flow of the base material during friction stir welding. Moreover, they indicated that homogenous structure could be achieved with a worm pin [31].

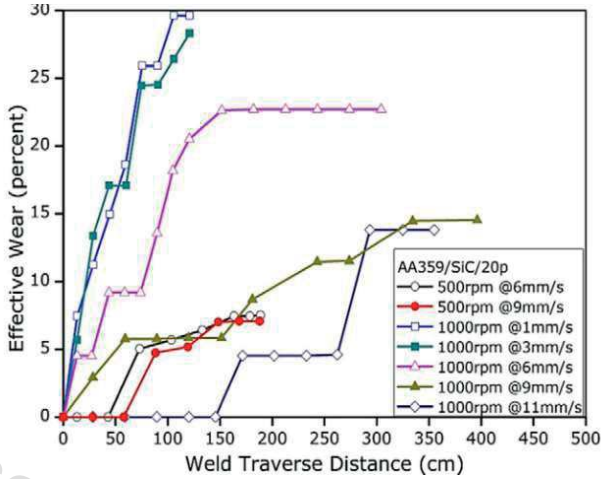


Fig. 7. Pin tool wear as a per cent of initial tool shape projections versus weld traverse [29]

2.4. Fatigue life evaluation

Friction stir processing improved the fatigue life of aluminium alloy Al-7Si-0.6Mg. FSP eliminate the porosity and refined the Si particles. The crack growth rate dropped due to elimination of the notch effect together with increased crack path tortuosity. The particles/matrix interface was responsible for crack nucleation in the FSP condition, size ratio analysis of the respective defect matches will with the Paris crack growth equation. As oscillatory crack growth behaviour was seen up to 450 µm for the cast condition, whereas in the FSP condition this behaviour stopped at 180 µm. The crack growth rate was found to be one order of magnitude higher in the cast condition than the FSP condition. Short crack behaviour was observed in both condition and the critical transition crack length matched will with the respective micro-structural characteristics dimension. The cast fatigue specimen showed a life of 45500 cycles. As expected crack were noted to have originated at porosity when the test was stopped after the first 5000 cycles which implies a 10% or lower crack initiation periods [32]. Cast A356 alloys prepared by FSP were subjected to fatigue investigated. Fatigue life improvement was attributed to significant refinement, homogenization of the microstructure and the elimination of porosity. FSP resulted in a significant breakup and uniform distribution of Si particles in the aluminium matrix as well as elimination of porosity. This lead to an improvement in fatigue stress threshold stress > 80% in the stir zone of the FSP sample. This improvement in fatigue properties is attributed to an increased effect crack length and reduction in crack growth rates. FSP can be used as a tool to locally modify the micro structures in

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

regions experiencing high fatigue loading and thus significant improve the overall performance of aluminium casting [33]. The friction stir welding (FSW) process, has been successfully applied to several aluminium alloys, leading to joint properties (such as tensile strength, toughness, and fatigue life) in some cases higher than those of the base material [34-40]. Low cycle fatigue life of the friction stir welding composite was always lower than that of the base material, high values of total strain amplitude were characterized by a fatigue life ratio around 1:2. The lower fatigue life of the FSW material could be related to the microstructural and surface modification induced by the process, which also led to a reduction of hardness and tensile strength.

The effect of surface finishing was particularly high for the two specimens tested at the lowest strain amplitude values, which gave premature failures starting from the high geometrical discontinuity between the base material and the shoulder in the retreating side of the welding (Fig. 8 (a)) with increasing the strain amplitude this effect was of lower entity, but failure occurred in the retreating side welded zone (Fig. 8 (b)-(c)). The possible presence of residual stresses due to the welding process should not have any influence on fatigue life in strain controlled test [41].

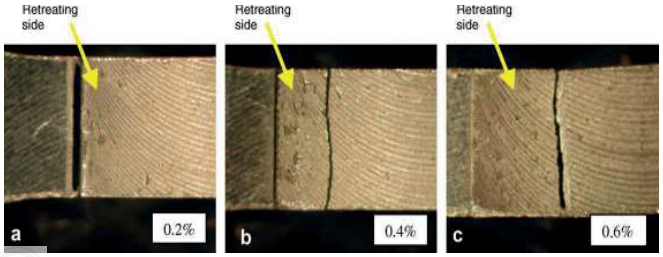


Fig. 8. Fracture appearance of FSW low-cycle fatigue specimens tested at different strain-amplitudes [41]

2.5. Effect of welding parameter

Recently many studies have been conducted to establish the optimum parameter for friction stir welding of dissimilar aluminium alloys and to identify their micro-structures, mechanical properties and defect formation. It is important to note that for FSW of dissimilar material, addition parameters, such as material arrangement and position of tool plunge with respect to the weld centre line, need to be considered, as well as general parameter such as tool geometry, rotation speed and welding speed. Because material flows and thermal hysteresis differ between the advancing and retreating sides, so material arrangement and tool plunge position exert a significant effect on weld formation. In friction stir welding process, the welding parameter including tool rotation speed, traverse and axial force affect the friction heat generation and mixing process. Therefore optimum welding parameter must be selected in order to produce the best joint strength. The efficiency of aluminium matrix composite weld joints is generally in the range of 60% to 97% of those of the base material. It is accepted that the ultimate tensile strength of friction stir welding joints of aluminium matrix composite increases by increasing the rotation speed until a specific limit [42-48].

The highest hardness value occurs in the centre of the nugget zone followed by a gradual decrease across the thermo-mechanically affected zone (TMAZ) and Heat affected zone (HAZ) until reaching the hardness value of the base material as shown in Figure 9. This is attributed to more grain refinement in the nugget zone due to dynamic recrystallization and more uniform distribution of finer reinforcement particles in the weld zone due to friction stir welding action. It may also be concluded that lower heat input leads to the formation of coarse grains because of incomplete recrystallization and thus a reduction of the nugget zone micro hardness [44].

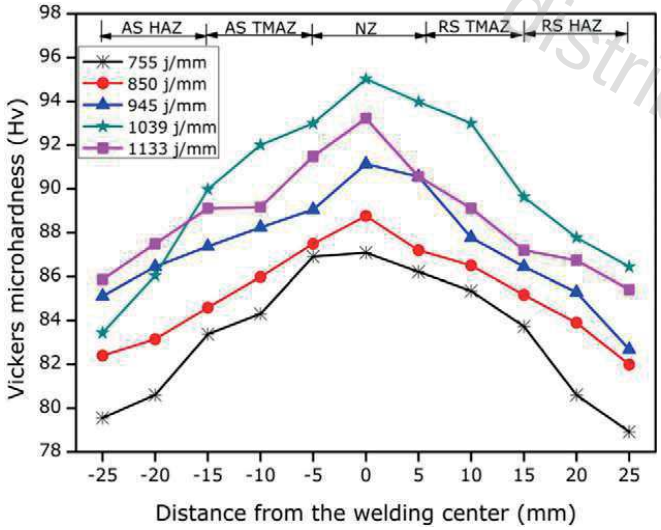


Fig. 9. Micro hardness profile across the weld region of AA6061/SiC/10p at different heat [44]

3. Conclusion

The mechanical properties of welded joint by friction stir welding are largely dependent on the combined effect of both the composition of alloying element and processing parameter, there for, the mechanical performance of friction stir welding joint should be evaluated accordingly. Past

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

researcher showed that FSW is a potential welding process to achieve defect free joint. Welding parameter such as tool rotation, transverse speed and axial force have a significant effect on the amount of heat generated and strength of FSW joints. Microstructure evaluation of FSW joints clearly shows the formation of new fine grains and refinement of reinforcement particles in the weld zone with different amount of heat input by controlling the welding parameter.

This review aims to outline the current working on friction stir welding with number of specific issue including welding parameter, different aluminium alloy, macro and microstructure, mechanical properties, tool wear and fatigue evaluation. Friction stir welding have a potential benefits in cost reduction, joint efficiency improvement and high production accuracy make it more attractive for non weldable series. The welding parameter such as traverse speed, tool rotation speed and axial force have a significant amount of heat generation and strength of friction stir welding joints. The microstructure calculation of FSW joints clearly shows the formation of new grains and refinement of reinforcement particles in the weld zone with different amount of heat input by controlling welding parameter. The welding parameter also affect the mechanical properties of FSW joints. The wear of FSW tools is a main issue when joining different material with the help of friction stir welding, the life of welding tool could be more, when we use correct processing parameter for different material and working conditions.

References

[1] M.B.D. Ellis, Joining of aluminiumbased metal matrix composites, International Materials Reviews 41/2 (1996) 41-58.

[2] R.Y. Huang, S.C. Chen, J.C. Huang, Electron and laser beam welding of high strain rate superplastic Al-6061/SiC composites, Metallurgical and Materials Transactions A 32/10 (2001) 2575-2584.

[3] T. Prater, Friction stir welding of metal matrix composites for use in aerospace structures, Acta Astronautica 93 (2014) 366-373.

[4] G. Lutjering, J.C Williams, Titanium, Springer-Verlag, New York, 2003.

[5] Y. Zhang, Y.S. Sato, H. Kokawa, S.H.C Park, S. Hrano, Stir zone microstructure of commercial purity titanium friction stir welded using pnBN ttol. Material Science and Engineering A 488 (2008) 25-30.

[6] J. Wang, J. Su, R.S. Mishra, R. Xu, J.A. Baumann, Tool wear mechanism in friction stir welding 321 (2014) 25-32.

[7] J.Q. Su, T.W. Nelson, R.S. Mishra, M. Mahoney, Microstructure investigation of friction stir welding 7050-T651 aluminum, Acta Materialia 51/3 (2003) 713-729.

[8] V.J. Arumoni, R.S. Mishra, Friction stir processing of Al-alloy for defence application, International Journal of Advance Research and Innovation 2/2 (2014) 337-341.

[9] K. Kolligan, Material flow behavior during friction stir welding of aluminum, Welding Research (1999) 229-237.

[10] G. Kohn, Y. Greenberg, I. Makover, A. Munitz, Laser Assisted, Friction Stir Welding, Welding Journal, American Welding Society (2002) 1-3.

[11] W.M. Thomas, Friction Stir Butt Welding International Patent Application, No. PCT/GB92 Patent Application No. 9125978.8, 1991.

[12] S.S. Yutaka, K. Hiroyuki, Microstructural evolution of 6063, aluminum during friction stir welding. Metallurgical and Materials Transactions A 30A (1999) 2429-2437.

[13] G. Liu, L.E. Murr, Microstructural aspects of the friction-stir welding of 6061-T6 aluminum, Scripta Materialia 37/3 (1997) 355-361.

[14] S.S. Yutaka et al. Microstructural factors governing hardness in friction-stir welds of solid-solution-hardened Al alloys. Metallurgical and Materials Transactions A 32A (2001) 3033-3041.

[15] K.W. Jata, S.L.Semiatin, Continuous dynamic recrystallization during friction stir welding of high strength aluminum alloys, Scripta Materialia 43 (2000) 743-749.

[16] S. Manoj, A. Navneet, Chandan Pandey, Husain Mehdi, Preliminary Studies on Thermal Cycling of Reactor Pressure Vessel Steel, International Journal of Mechanical Engineering 4/2 (2014) 51-58.

[17] S. Manoj, A. Navneet, P. Chandan M. Husain, Mechanical Properties of Bimetallic Weld Joint between SA 516 Grade 65 Carbon Steel and SS 304 L for Steam Generator Application, International Journal of Research in Engineering and Technology 3/7 (2014) 39-42.

[18] Z.Y. Ma, R.S. Mishra, M.W. Mahoney, R. Grimes, High srain rate super plasticity in friction stir processed Al-4Mg-1Zr alloy, Material Science and Engineering A 351 (2003) 148-153.

[19] Y. Wang, X.L. Shi, R.S. Mishra, T.J Watson, Friction stir welding of devitrified AL-4.0Y-4.0Ni-0.9Co alloy produced by amorphose powder, Scripta Materialia 56 (2007) 971-974.

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

[20] R.S. Mishra, M.W. Mahoney, S.X. McFadden, N.A. Mara, A.K. Mukherjee, High strain rate superplasticity in a friction stir processed 7075 Al alloy Scripta Materialia 42 (2000) 163-168.
[21] R.S. Mishra, M.W. Mahoney, Friction Stir Processing: A New Grain Refinement Technique to Achieve High Strain Rate Superplasticity in Commercial Alloys Materials Science Forum 357-359 (2001) 507-514.
[22] R.S. Mishra, M.W. Mahoney, S.X. McFadden, N.A. Mara, A.K. Mukherjee, High strain rate superplasticity in a friction stir processed 7075 Al alloy Scripta Materialia 42 (2000) 163-168.
[23] I. Charit, R.S. Mishra, Evaluation of microstructure and superplasticity in friction stir processed 5083 Al alloy, Journal of Materials Research 19/4 (2004) 3329-3342.
[24] I. Charit, R.S. Mishra, High strain rate superplasticity in a commercial 2024 Al alloy via friction stir processing, Materials Science and Engineering: A 359 (2003) 290-296.
[25] I. Charit, R.S. Mishra, Low temperature superplasticity in a friction-stir-processed ultrafine grained Al-Zn-Mg-Sc alloy, Acta Materialia 53 (2005) 4211-4223.
[26] R.S. Mishra, Z.Y. Ma, I. Charit, Friction stir processing: A Novel technique for fabrication of surface composite, Matrial Science and Engineering A 341 (2003) 307-310.
[27] J.Q. Su, T.W. Nelson, R.S. Mishra, M. Mahoney, Microstructure investigation of friction stir welding 7050-T651 aluminum Acta Materialia 53/1 (2003) 713-729.
[28] S. Jana, R.S. Mishra, J.A Baumann, G. Grant, Effect of process parameters on abnormal grain growth during FSP of a cast Al-Alloy, Material Science and Engineering A 528 (2010) 189-199.
[29] R. Rai, A. De, H.K.D.H. Bhadeshia, T. DebRoy, Review: friction stir welding tools, Science and Technology of Welding and Joining 16/4 (2011) 325-342.
[30] R.S. Mishra, Z.Y. Ma, Friction stir welding and processing, Materials Science and Engineering: R: Reports 50/1-2 (2005) 1-78.
[31] J. Wang, J. Su, R.S. Mishra, R. Xu, J.A. Baumann, Tool wear mechanism in friction stir welding, Wear 321 (2014) 25-32.
[32] G.J. Fernandez, L.E. Murr, Characterization of tool wear and weld optimization in the friction-stir welding of cast aluminum 359 + 20% SiC metal-matrix composite, Materials Characterization 52/1 (2004) 65-75.
[33] S. Jana, R.S. Mishra, J.B. Baumann, Effect of friction stir processing on fatigue behavior of an investigation cast Al-7Si-0.6Mg alloy, Acta Materialia 58 (2010) 989-1003.
[34] S.R. Sharma, R.S. Mishra, Z.Y. Ma, Effect of friction stir processing on fatigue behaviour of A356 Alloy, Scripta Materialia 51 (2004) 237-241.
[35] C.G. Rhodes, M.W. Mahoney, W.H. Bingel, R.A. Spurling, C.C. Bampton, Effects of friction stir welding on microstructure of 7075 aluminium, Scripta Materialia 36/1 (1997) 69-75.
[36] K.V. Jata, S.L. Semiatin, Continuous dynamic recrystallization during friction stir welding of high strength aluminum alloys, Scripta Materialia 43 (2000) 743-749.
[37] A.J. Leonard, Microstructure and ageing behaviour of FSWs in aluminium alloys 2014T651 and 7075-T651. In: Proc II Friction Stir Welding Symposium. Goteborg, Sweden, 2000, 26-28.
[38] G. Liu, L.E. Murr, C.S. Niou, J.C. McCloure, F.R. Vega, Microstructural aspects of the friction-stir welding of 6061-T6 aluminium, Scripta Materialia 37/3 (1997) 355-361.
[39] W.B. Lee, Y.M. Yeon, S.B. Jung, The improvement of mechanical properties of friction-stir-welded A356 Al alloy, Matrial Science and Engineering: A 355 (2003) 154-159.
[40] M.P. Miles, B.J. Decker, T.W. Nelson, Formability and Strength of Friction-Stir-Welded Aluminum Sheets, Metallurgical and Materials Transactions: A 35 (2004) 3461-3468.
[41] M.J. Jones, L.P. Heurtier, C .Desrayaud, F. Montheillet, D. Allehaux, J.H..Driver, Correlation between microstructure and microhardness in a friction stir welded 2024 aluminium alloy, Scripta Materialia 52/8 (2005) 693-697.
[42] L. Ceschini, I. Boromei, G. Minak, A. Morri, F. Tarterini, Microstructure, tensile and fatigue properties of AA6061/20 vol.%Al_2O_{3p} friction stir welded joints, Composites: Part A 38 (2007) 1200-1210.
[43] H. Nami, H. Adgi, M. Sharifitabar, H. Shamabadi, Microstructure and mechanical properties of friction stir welded Al/Mg2Si metal matrix cast composite, Materials and Design 32/2 (2011) 976-983.
[44] N. Murugan, B. Ashok Kumar, Prediction of tensile strength of friction stir welded stir cast AA6061-T6/AlNp composite, Materials and Design 51 (2013) 998-1007.
[45] P. Periyasamy, B. Mohan, V. Balasubramanian, Effect of heat input on mechanical and metallurgical

Thiscopy is for personal use only - distributi

- This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited. - This copy is for personal use only - distribution prohibited.

properties of friction stir welded AA6061-10% SiCp MMCs, Journal of Materials Engineering and Performance 21/11 (2012) 2417-2428.
[46] P. Periyasamy, B. Mohan, V. Balasubramanian, S. Rajakumar, S. Venugopal, Multi objective optimization of friction stir welding parameters using desirability approach to join Al/SiCp metal matrix composites, Transactions of Nonferrous Metals Society of China 23/4 (2013) 942-955.
[47] I. Dinaharan, N. Murugan, Optimization of friction stir welding process to maximize tensile strength of AA6061/ZrB2 in-situ composite butt joints, Metals and Materials International 18/1 (2012) 135-142.
[48] K. Kalaiselvan, N. Murugan, Role of friction stir welding parameters on tensile strength of AA6061-B4C composite joints, Transactions of Nonferrous Metals Society of China 23/3 (2013) 616-624.
[49] Y. Bozkurt, A. KentlI, H. Uzun, S. Salman, Experimental investigation and prediction of mechanical properties of friction stir welded aluminium metal matrix composite plates, Materials Science 18/4 (2012) 336-340.

This copy is for personal use only - distributi

IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Mechanisms of plasma-assisted catalyzed growth of carbon nanofibres: a theoretical modeling

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

2017 Plasma Sources Sci. Technol. 26 024006

(http://iopscience.iop.org/0963-0252/26/2/024006)

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 01/02/2017 at 09:49

Please note that terms and conditions apply.

You may also be interested in:

Formation of vertically aligned carbon nanostructures in plasmas
I Denysenko and N A Azarenkov

Plasma heating effects in catalyzed growth of carbon nanofibres
I Denysenko and K Ostrikov

Low-and high-temperature controls in carbon nanofiber growth in reactive plasmas
H Mehdipour, K Ostrikov and A E Rider

Investigating the plasma chemistry for the synthesis of carbon nanotubes/nanofibres in an inductively coupled plasma enhanced CVD system
M Mao and A Bogaerts

Challenges in process integration of catalytic DC plasma synthesis of vertically aligned carbon nanofibres
Anatoli V Melechko, Ryan C Pearce, Dale K Hensley et al.

Computer modelling of the plasma chemistry and plasma-based growth mechanisms for nanostructured materials
Annemie Bogaerts, Maxie Eckert, Ming Mao 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.

Plasma nanotechnology: past, present and future
M Meyyappan

IOP Publishing Plasma Sources Science and Technology
Plasma Sources Sci. Technol. **26** (2017) 024006 (10pp) doi:10.1088/1361-6595/aa5120

Mechanisms of plasma-assisted catalyzed growth of carbon nanofibres: a theoretical modeling

R Gupta, S C Sharma and R Sharma

Department of Applied Physics, Delhi Technological University (DTU), Shahbad Daulatpur, Bawana Road, Delhi-110042, India

E-mail: suresh321sharma@gmail.com

Received 10 June 2016, revised 28 November 2016
Accepted for publication 1 December 2016
Published 20 January 2017



Abstract

A theoretical model is developed to study the nucleation and catalytic growth of carbon nanofibers (CNFs) in a plasma environment. The model includes the charging of CNFs, the kinetics of the plasma species (neutrals, ions and electrons), plasma pretreatment of the catalyst film, and various processes unique to a plasma-exposed catalyst surface such as adsorption of neutrals, thermal dissociation of neutrals, ion induced dissociation, interaction between neutral species, stress exerted by the growing graphene layers and the growth of CNFs. Numerical calculations are carried out for typical glow discharge plasma parameters. It is found that the growth rate of CNFs decreases with the catalyst nanoparticle size. In addition, the effect of hydrogen on the catalyst nanoparticle size, CNF tip diameter, CNF growth rate, and the tilt angle of the graphene layers to the fiber axis are investigated. Moreover, it is also found that the length of CNFs increases with hydrocarbon number density. Our theoretical findings are in good agreement with experimental observations and can be extended to enhance the field emission characteristics of CNFs.

Keywords: PECVD, kinetics of plasma species, carbon nanofiber, catalyst nanoparticle

(Some figures may appear in colour only in the online journal)

1. Introduction

The growth of carbon nanofibers (CNFs), carbon nanotubes (CNTs) and related carbon nanostructures have gained heightened research and commercial interest in the last few decades. They can be produced by many pathways such as laser ablation [1], arc discharge [2], thermal chemical vapor deposition (CVD) [3], and plasma enhanced chemical vapor deposition (PECVD). Amidst them, the PECVD technique is exclusively used to synthesize vertically oriented carbon nanostructures. The incited electric field of the plasma sheath applies force on the carbon nanostructures and directs their growth in a vertical direction [4, 5]. It has been observed that nanostructures produced via plasma-based processes exhibit various miraculous electrical, physical and optical properties, and, as a result, application in electronics, in particular field emission devices [6–8], mechanics, biochemistry and several others [9].

CNFs and CNTs are graphitic filaments having a cylindrical shape with radii on the nanometer scale and lengths up to several micrometers. Although, CNFs/CNTs can be oriented from the same catalyst and have similar external morphology, they differ in terms of diameter and crystalline structure. The orientation of their graphite basal planes with respect to the growth axis characterizes the divergence between CNFs and CNTs, and dictates their properties. CNFs are composed of graphitic funnels whereas concentric cylinders exist in CNTs [10]. In general, graphitic planes are aligned along the growth axis in CNTs, i.e. inclination angle $\theta = 0$, and are inclined at some non-zero angle to the growth axis in the case of CNFs. Diameter, height, shape and orientation of the vertically aligned CNFs (VACNFs) must be externally controlled for various practical applications [11, 12]. The low temperature PECVD technique favors the disciplined growth of VACNFs

 © 2017 IOP Publishing Ltd

as several parameters, such as ion, neutral and electron fluxes from the plasma, plasma power, gas mixture ratio in the plasma, and temperature have an intense effect on VACNF growth [10, 13–15]. In addition to growth, the plasma also affects the nucleation of the CNF. The metal catalyst nanoparticles, which act as the nucleation sites for the growth of CNFs, play an important role during CNF growth and can be formed by plasma etching or plasma processing of the deposited (on the substrate surface) metal catalyst film prior to CNF growth [16]. It was observed that the size of the catalyst nanoparticle governs the diameter, shape, alignment and crystalline structure of the CNFs [17, 18]. The initial size of the catalyst nanoparticle can be tuned by varying plasma conditions [19, 20].

In the present paper, we attempt to provide an in-depth study of nucleation and the deterministic growth of VACNFs, accounting for plasma sheath effects, kinetics of plasma species (electrons, ions, and neutrals), formation of catalyst nanoparticles from catalyst film, and the various other processes involved on the surface of the catalyst particle during the growth of CNFs in the plasma environment. In particular, we solve the simultaneous coupled kinetic equations to analyze the effect of plasma species on the size of the catalyst nanoparticles and the growth of CNFs. In section 2, a theoretical model is introduced, and results and discussion are given in section 3. Finally, a conclusion is given in section 4.

2. Model

The present model considers the following scenario towards elucidating the nucleation and the growth mechanism of CNFs in the plasma.

I. The pre-etching of the catalyst film in the presence of plasma segregates the catalyst film into discrete islands (nanoparticles) [21]. These catalyst nanoparticles act as the fundamental sites for the nucleation and growth of CNFs. The initial geometry of the catalyst nanoparticles can greatly affect the configuration of the resulting carbon nanostructure [18].
II. Carbon precursors, i.e. hydrocarbon gases and hydrogen are decomposed and adsorbed via various processes on the catalyst nanoparticle surface to produce carbon species and active hydrogen radicals. These carbon species dissolve and diffuse (surface as well as bulk) through catalyst nanoparticles and ultimately precipitate at the rear of the catalyst nanoparticle to form CNFs [22, 23].
III. Following this carbon precipitation, graphene layers form around the catalyst nanoparticle. These growing graphene layers exert significant stress on the catalyst particle resulting in the deformation of its shape [24].
IV. The CNF grows vertically due to alignment forces induced by the electric field provided by the plasma sheath [5].

In addition to these elementary processes, we developed a theoretical model considering the reactive plasma of a $Ar + H_2 + C_2H_2$ gas mixture containing electrons, positively charged ions of acetylene ($C_2H_2^+$) and hydrogen (H^+) denoted as ions of type 1 and 2, respectively, neutrals of acetylene C_2H_2 (type 1) and hydrogen H (type 2), and a nickel (Ni) catalyst film placed over a silicon (Si) substrate surface and the sheath configuration near the substrate surface. The electric field within the sheath is assumed to be along the z-direction. Following Tewari and Sharma [25], Mehdipour *et al* [26], and Lieberman *et al* [27], the equations to investigate the sheath kinetics are

continuity equation:

$$\left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y} + \hat{k}\frac{\partial}{\partial z}\right)(n_{0k}\overrightarrow{\nu_{0k}}) = f_{0i}\, n_{0e}, \tag{1}$$

Ion−Momentum balance equation:

$$M_{0k} n_{0k} v_{0k} \frac{dv_{0k}}{dz} = e n_{0k} E - M_{0k} n_{0k} f_{0kn} v_{0k}, \tag{2}$$

$$\text{Poisson equation: } \frac{\mathrm{d}^2\phi(x)}{\mathrm{d}z^2} = -4\pi\sum q_{0k}\Omega_{0k} n_{0k}, \tag{3}$$

where k refers to either electron $C_2H_2^+$or H^+, n_{0k} is their number density, M_{0k} is their mass in the plasma, ν_{0k} is their fluid velocity, f_{0i} is the ionization frequency, E is the electric field, f_{0kn} is the collision frequency, Ω_{0k} is the kth ion-to-electron number density ratio, q_{0k} is the charge of the species, ϕ is the electrostatic potential, $\sum_k \Omega_{0k} = 1$, and $0 < \Omega_{0k} < 1$.

2.1. Charging the CNFs

Equation (4) describes the charge developed on the CNF surface due to accretion of positively charged ions and electrons on the surface of the CNF. The effects of collisional processes and dissipative losses have not been accounted for in this equation since plasmas are considered here to be thermodynamically open systems [28, 29]:

$$\dot{q} = \sum_{j=1}^{2}(I_{ij})_{\text{tip}} + \sum_{j=1}^{2}(I_{ij})_{\text{cur}} - \gamma_e[(I_e)_{\text{tip}} + (I_e)_{\text{cur}}] \tag{4}$$

where q is the charge over the whole CNF including the tip as well as the curved surface, which can be expressed as $q = Ze$, where Z is the charge number and e is the electronic charge. For negatively charged nanoparticles/nanostructures, Z may vary from -1 to -100.

$(I_{ij})_{\text{tip}}$ and $(I_{ij})_{\text{cur}}$ are the ion collection currents at the tip and curved surface of the CNF and given by,

$$(I_{ij})_{\text{tip}} = \frac{\pi d^2}{4}\left(\frac{8k_B T_{ij}}{\pi m_{ij}}\right)^{1/2} n_{ij}(x)[1 - q\alpha_i]\exp \times \left[\frac{-E_b}{k_B T_S}\right]\exp\left[\frac{-eU_S}{k_B T_S}\right], \tag{5}$$

$$(I_{ij})_{cur} = n_{ij}(x)\frac{dh}{2}\left(\frac{2\pi k_B T_{ij}}{m_{ij}}\right)^{1/2} \times \left\{\frac{2}{\sqrt{\pi}}\left(\frac{eV_S}{k_B T_{ij}}\right)^{1/2} + \exp\left[\frac{eV_S}{k_B T_{ij}}\right]\mathrm{erfc}\left[\left(\frac{eV_S}{k_B T_{ij}}\right)^{\frac{1}{2}}\right]\right\} \times \exp\left[\frac{-E_b}{k_B T_S}\right]\exp\left[\frac{-eU_S}{k_B T_S}\right], \quad (6)$$

where d is the diameter of the CNF tip, h is the height of the CNF, T_{ij} is the ion temperature, m_{ij} is the mass of ion (subscript j refers to type 1 or 2 of positively charged ion species as explained earlier), k_B is the Boltzmann constant, $\alpha_i = \left(\frac{2e^2}{dk_B T_S}\right)$, V_S is the surface potential of the curved surface of the CNF, U_S is the substrate bias, E_b is the energy barrier for bulk diffusion, T_S is the substrate or catalyst nanoparticle temperature, $n_{ij}(x) = n_{ij0}\left(1 - \frac{2e\phi(x)}{m_{ij}v_{i0}^2}\right)$ is the ion density in the plasma sheath where $\phi(x)$ is the electrostatic potential, n_{ij0} is the initial ion density in plasma sheath, v_{i0} is the ion velocity at any point within the plasma and $\phi(x) = \phi_0 \exp\left(-\frac{|x|}{\lambda_d}\right)$, where ϕ_0 is the negative potential at the surface and $\lambda_d = \sqrt{\frac{k_B T_e}{n_e e^2}}$ is the Debye length. $(I_e)_{tip}$ and $(I_e)_{cur}$ are the electron collection currents at the tip and curved surface of the CNF and given by,

$$(I_e)_{tip} = \pi\frac{d^2}{4}\left(\frac{8k_B T_e}{\pi m_e}\right)^{1/2} n_e(x)\exp\left[q\alpha_e + \frac{eU_S}{k_B T_S}\right] \quad (7)$$

$$(I_e)_{cur} = n_e(x)\frac{d}{2}h\left(\frac{2\pi k_B T_e}{m_e}\right)^{1/2}\exp\left[\frac{eV_S}{k_B T_e} + \frac{eU_S}{k_B T_S}\right] \quad (8)$$

where $n_e(x) = n_{e0}\exp\left[\frac{|e|\phi(x)}{k_B T_e}\right]$ is the electron density in the plasma sheath, n_{e0} is the initial electron density in the plasma sheath, T_e is the electron temperature, m_e is the mass of an electron, $\alpha_e = \left(\frac{2e^2}{dk_B T_e}\right)$, and γ_e is the sticking coefficient of electrons, which describes the probability of electrons adsorption to a surface, i.e. the ratio of the number of electrons that adsorb to a surface to the number of electrons impinge upon the surface [30].

2.2. Particle balance equations for different plasma species

$$\dot{n}_e = \sum_{j=1}^{2}\beta_j n_j - \sum_{j=1}^{2}\alpha_j n_e n_{ij} - \gamma_e n_{cnf}[(I_e)_{tip} + (I_e)_{cur}], \quad (9)$$

$$n_{i1} = \beta_1 n_1 - \alpha_1 n_e n_{i1} - n_{cnf}[(I_{i1})_{tip} + (I_{i1})_{cur}] - J_{ai1} + J_{desp1}, \quad (10)$$

$$\dot{n}_{i2} = \beta_2 n_2 - \alpha_2 n_e n_{i2} - n_{cnf}[(I_{i2})_{tip} + (I_{i2})_{cur}] - J_{ai2} + J_{desp2} + J_{thd}, \quad (11)$$

$$\dot{n}_j = \alpha_j n_e n_{ij} - \beta_j n_j + n_{cnf}(1 - \gamma_{ij})[(I_{ij})_{tip} + (I_{ij})_{cur}] - n_{cnf}\gamma_j[(I_j)_{tip} + (I_j)_{cur}], \quad (12)$$

where β_j is the coefficient of ionization of the constituent neutral atoms due to an external field, n_j is the number density of the neutral atoms, n_e is the electron number density, n_{ij} is the ion number density, $\alpha_j(T_e) = \alpha_{j0}\left(\frac{300}{T_e}\right)^k$ cm^3 s^{-1} is the coefficient of recombination of electrons and positively charged ions, $k = -1.2$ is a constant, γ_{ij} is the ion sticking coefficient, γ_j is the sticking coefficient of neutral atoms, and j refers to either 1 or 2 type of positively charged ions and neutrals. $J_{aij} = \frac{p_i}{(2\pi m_{ij}k_B T_{ij})^{1/2}} \times \frac{n_{ij}}{j_{ij}}$ is the adsorption flux onto the catalyst–substrate surface, p_i is the partial pressure of adsorbing species, $J_{desp} = j_{ij}\nu\exp\left(-\frac{\varepsilon_{ai}}{k_B T_{ij}}\right)$ is the desorption flux from the catalyst–substrate surface, ε_{ai} is the adsorption energy, j_{ij} is the ion flux on the catalyst–substrate surface, ν is the thermal vibrational frequency, $J_{thd} = j_H\nu\exp\left(\frac{-\delta\varepsilon_{th}}{k_B T_S}\right)$ is the flux of type 2 ions (namely hydrogen) on account of thermal dehydrogenation, $\delta\varepsilon_{th}$ is the activation energy of thermal dehydrogenation, j_H is the hydrogen ion flux at the catalyst–substrate surface [25]. $(I_j)_{tip}$ and $(I_j)_{cur}$ are the neutral collection current at the tip and the curved surface of the CNF and are given by:

$$(I_j)_{tip} = n_j(x)\pi\frac{d^2}{4}\left(\frac{8k_B T_j}{\pi m_j}\right)^{1/2}, \quad (13)$$

$$(I_j)_{cur} = n_j(x)\pi\frac{d}{2}h\left(\frac{2\pi k_B T_j}{m_j}\right)^{1/2}, \quad (14)$$

where T_j is the neutral atom temperature and m_j is the mass of the neutral atom.

Equation (9) indicates the balance of the electron number density in the plasma system on account of ionization of neutral atoms, recombination of electrons and ions, and electron collection current on the CNF surface. The first term on the right-hand side of equation (9) refers to the gain in electron density per unit time due to ionization of neutral atoms; the second and third terms represent the rate of decrease in electron density due to electron–ion recombination and the electron collection current at the CNF surface, respectively.

Equations (10) and (11) describe the balance of positively charged ion species in the plasma system due to ionization of neutral atoms, electron–ion recombination, ion collection current at the surface of the CNF, their adsorption, desorption and thermal dehydrogenation. The first term on the right-hand side describes the gain in ion density per unit time on account of ionization of neutral atoms. The second and third terms represent the rate of decrease in positively charged ion density due to electron–ion recombination and the ion collection current at the surface of the CNF (tip as well as the curved surface), respectively. The fourth and fifth terms indicate adsorption and desorption of ions to/from the catalyst–substrate surface, respectively.

One additional term in equation (11), i.e. the last term, represents the increase in the hydrogen ion number density in the plasma because of thermal dehydrogenation of hydrocarbon ions, since at temperatures of >400 K the dehydrogenation process (removal of hydrogen) possibly becomes an important factor [31, 32].

Equation (12) is the particle balance equation for neutrals of type 1 and 2, which indicates the growth rate of the neutral atom number density. The first and third terms show the increase in growth rate of the neutral density due to electron–ion recombination and neutralization of ions collected at the surface of the CNF, respectively. The second and fourth terms symbolize the decay rate of neutral density due to ionization and accumulation of neutrals on the surface of the CNF, respectively.

2.3. Formation of catalyst nanoparticles

Thermal and energy balance play a significant role in the plasma processing of solid surfaces. Here, plasma processing refers to the etching process in the presence of plasma that is responsible for segregation of thin film into nanoparticles. In this process, the surface is heated due to energy transfer. Therefore, the general power balance at the surface of the catalyst film placed over the silicon substrate in the presence of plasma, i.e. hydrogen plasma is given by the following equation [33].

$$P_{\text{input}} = P_{\text{gain}} + P_{\text{loss}}, \tag{15}$$

where P_{input}, P_{gain} and P_{loss} are the total input plasma power, power gain and power loss during plasma processing, respectively. The total input plasma power is the contribution of various energy fluxes per unit time.

$$\begin{aligned} P_{\text{input}} &= A_{\text{f}}\sigma\left[\in_{\text{env}}(T_{\text{env}})^4 - \in_{\text{f}}(T_{\text{S}})^4\right] \\ &+ \left[I_{\text{ef}}\varepsilon_{\text{e}}^{\text{f}} + I_{\text{i2f}}\varepsilon_{\text{i2}}^{\text{f}} + I_{\text{2f}}\varepsilon_{2}^{\text{f}}\right] \\ &+ \frac{3}{2}k_{\text{B}}T_{\text{S}}(1-\gamma_{\text{i2}})I_{\text{i2f}}, \end{aligned} \tag{16}$$

where A_{f} is the area of the metal catalyst film, σ (=5.67 × 10^{-5} ergs $\text{cm}^{-2}\,\text{s}^{-1}\,\text{K}^{-4}$) is Stefan's constant, $\in_{\text{env}}$ (≈1) is the emissivity of the environment, $\in_{\text{f}}$ (=0.12 for nickel) is the emissivity of the material of the metal catalyst film, T_{S} is the substrate temperature, T_{env} is the temperature of the plasma environment, $\varepsilon_{\text{e}}^{\text{f}}$, $\varepsilon_{\text{i2}}^{\text{f}}$, and $\varepsilon_{2}^{\text{f}}$ are the mean energies collected by the metal catalyst film surface due to electrons, ions and neutral atoms, respectively [34], and are given by $\varepsilon_{k}^{\text{f}} = \left[\left(\frac{2-Z\gamma_k}{1-Z\gamma_k}\right) - Z\gamma_k\right]k_{\text{B}}T_k$ and $\gamma_k = \frac{E_{\text{b}}}{k_{\text{B}}T_k}$, where k can be an an ion, electron or neutral atom, E_{b} is the energy barrier at the catalyst film surface, I_{i2f}, I_{ef}, and I_{2f} are the collection currents at the metal catalyst film surface due to ions of type 2, electrons, and neutrals of type 2, respectively [35], and are given by

$$\begin{aligned} I_{\text{i2f}} &= n_{\text{i2}}(x)A_{\text{f}}\left(\frac{k_{\text{B}}T_{\text{i2}}}{2\pi^2 m_{\text{i2}}}\right)^{1/2}\left\{\frac{2}{\sqrt{\pi}}\left(\frac{eV_{\text{S}}}{k_{\text{B}}T_{\text{i2}}}\right)^{1/2}\right. \\ &\left. + \exp\left[\frac{eV_{\text{S}}}{k_{\text{B}}T_{\text{i2}}}\right]\text{erfc}\left[\left(\frac{eV_{\text{S}}}{k_{\text{B}}T_{\text{i2}}}\right)^{1/2}\right]\right\} \\ &\times \exp\left[-\frac{eU_{\text{S}}}{k_{\text{B}}T_{\text{S}}}\right], \end{aligned} \tag{17}$$

$$I_{\text{ef}} = \left(\frac{k_{\text{B}}T_{\text{e}}}{2\pi^2 m_{\text{e}}}\right)^{1/2} n_{\text{e}}(x)(A_{\text{f}})\exp\left[\frac{eV_{\text{S}}}{k_{\text{B}}T_{\text{e}}} + \frac{eU_{\text{S}}}{k_{\text{B}}T_{\text{S}}}\right], \tag{18}$$

$$I_{\text{2f}} = n_2 A_{\text{f}}\left(\frac{k_{\text{B}}T_2}{2\pi^2 m_2}\right)^{1/2}. \tag{19}$$

The first term in equation (16) denotes the heat radiation per unit time towards the surface of the film from the surroundings [25]. The second term represents the energy per unit time collected by the film surface due to electrons, ions and neutral atoms, and the last term is the energy gained per unit time due to formation of neutrals at the film surface.

$$\begin{aligned} P_{\text{gain}} &= \frac{\text{d}}{\text{d}t}(m_{\text{f}}C_{\text{f}}T_{\text{S}}) \\ &= A_{\text{f}}C_{\text{f}}d_{\text{f}}\rho_{\text{f}}\left[\frac{\text{d}T_s}{\text{d}t} + n_{\text{a}}\frac{\text{d}}{\text{d}t}\left(\frac{1}{6}\pi D^3\right)\right] \end{aligned} \tag{20}$$

where m_{f} is the mass of the catalyst film, d_{f} is the thickness of the film, C_{f} (=0.104 cal $\text{g}^{-1}\,{}^{0}\text{C}^{-1}$) is the specific heat of the film, ρ_{f} is the density of the catalyst film material, D is the diameter of the catalyst nanoparticle and n_{a} is the number of catalyst nanoparticles per unit area. Equation (20) represents the power gained when the catalyst film is etched in the presence of plasma. The applied plasma power fragments the catalyst film into the catalyst nanoparticles [21], and during this plasma treatment the temperature of the substrate increases and attains a constant value after some time [20]. Similarly to the power input, various processes contribute to the power loss (power out) during plasma pretreatment of the catalyst film:

$$\begin{aligned} P_{\text{loss}} &= \sigma A_{\text{f}}(\in_{\text{f}}T_{\text{S}}^4 - \in_{\text{env}}T_{\text{env}}^4) \\ &+ \delta_{\text{ef}}(1-\gamma_{\text{e}})\left(\varepsilon_{\text{ef}}^{\text{S}} - \frac{3}{2}k_{\text{B}}T_{\text{S}}\right)I_{\text{ef}} \\ &+ \left[\upsilon_0\sigma_{\text{ads}}J_2U_{\text{B}} + J_{\text{i2}}y_{\text{d}}(E_{\text{i}})(1-\theta_{\text{t}})\right]A_{\text{f}}, \end{aligned} \tag{21}$$

where $\delta_{\text{ef}} = 2(m_{\text{e}}/m_{\text{f}})$ is the fraction of energy lost during collision with the film, $\varepsilon_{\text{ef}}^{\text{S}} = 2k_{\text{B}}T_{\text{e}}$ is the mean energy of electrons collected by the surface film [34], υ_0 is the number of sites available, σ_{ads} (=6.8 × 10^{-16} cm^2) is the cross section for the reaction of atomic hydrogen [31], $J_j = n_j(k_{\text{B}}T_j/m_j)^{1/2}$ is the neutral atom flux, $J_{\text{ij}} = n_{\text{ij}}(k_{\text{B}}T_{\text{ij}}/m_{\text{ij}})^{1/2}$ is the hydrogen ion flux [31] (here j is 2 for the hydrogen atom), U_{B} (=3.74 eV for nickel) is the surface binding energy of the material (sputter target), y_{d} is the sputtering yield [43], and θ_{t} is the total surface coverage.

The first term of equation (21) represents the radiative energy per unit time from the surface to the surroundings. The

remaining terms denote the energy loss due to sticking and elastic collision of electrons with the surface, energy loss due to chemical etching, and energy loss due to sputtering [36], respectively. The resulting value of the diameter of the catalyst nanoparticles at time τ is assumed to be the outer diameter of the CNF (D).

2.4. Growth rate equation of carbon species on the catalyst surface

$$\dot{n}_{\mathrm{C}} = J_1(1-\theta_{\mathrm{t}}) + 2n_{\mathrm{C_2H_2}}\upsilon\exp\left(\frac{-\delta E_{\mathrm{i}}}{k_{\mathrm{B}}T_{\mathrm{S}}}\right) + \frac{2n_{\mathrm{C_2H_2}}J_{\mathrm{i1}}y_{\mathrm{d}}}{\upsilon_0} + \frac{J_{\mathrm{C}}}{\pi D} - \frac{J_2\sigma_{\mathrm{ads}}J_{\mathrm{H}}}{\upsilon} - n_{\mathrm{C_2H_2}}\upsilon\exp\left(\frac{-\delta E_{\mathrm{i}}}{k_{\mathrm{B}}T_{\mathrm{S}}}\right) - \frac{J_{\mathrm{i1}}\sigma_{\mathrm{ads}}J_{\mathrm{i2}}}{\upsilon}, \tag{22}$$

where n_{C} is the number of carbon species per unit area generated on the catalyst surface, $J_\alpha = (n_\alpha v_{\mathrm{th}\alpha}/4)$ is the flux of impinging species (where α = C, H) [31], $n_{\mathrm{C_2H_2}} = \theta_{\mathrm{C_2H_2}}\upsilon_0$ is the number density per unit area of C_2H_2 radicals [31], υ_0 ($=1.3\times10^{15}\,\mathrm{cm^{-2}}$) is the number of adsorption sites per unit area [31], and δE_{i} ($=1.87$ eV) is the thermal energy barrier.

Equation (22) describes the generation of carbon species on the catalyst surface via various complex processes such as adsorption of hydrocarbons (first term), thermal dissociation of hydrocarbons (second term), ion induced dissociation of hydrocarbons (third term), and direct carbon flux towards the catalyst surface (fourth term) [31]. The equations also include the processes that are responsible for carbon loss from the catalyst surface. These processes are: loss of carbon due to interaction of neutral hydrocarbons with neutral hydrogen (fifth term), carbon evaporation (sixth term), and ion–ion recombination (last term), i.e. hydrocarbon ion and hydrogen ion recombination [31].

2.5. Growth rate equation of CNF

$$\frac{\mathrm{d}}{\mathrm{d}t}[\pi(D^2-d^2)h] = \left[\frac{D_{\mathrm{S}}n_{\mathrm{C}}\times(\pi D^2)}{\upsilon\times(\pi D)} + \frac{D_{\mathrm{b}}n_{\mathrm{C}}\times\pi D}{2\upsilon} + 4kn_{\mathrm{C}}\times\pi D^2\right]\times\left(\frac{p}{\rho_{\mathrm{cnf}}I_{i1}}\right) + J_2(1-\theta_t) + J_2\exp\left(\frac{-\delta E_{\mathrm{th}}}{k_{\mathrm{B}}T_{\mathrm{S}}}\right) + \frac{n_{\mathrm{C_2H_2}}J_{i2}y_{\mathrm{d}}}{\upsilon_0} + n_{\mathrm{C_2H_2}}\nu\exp\left(\frac{-\delta E_i}{k_{\mathrm{B}}T_{\mathrm{S}}}\right) + J_{i2} - J_2\exp\left(\frac{-E_{a\mathrm{H}}}{k_{\mathrm{B}}T_{\mathrm{S}}}\right) + \frac{16D_{\mathrm{m}}m_{\mathrm{cat}}}{\rho_{\mathrm{cat}}\times\pi D^2}, \tag{23}$$

where d is the inner diameter of the CNF, h is the height of the CNF, $D_{\mathrm{S}} = D_{\mathrm{S0}}\exp(-E_{\mathrm{S}}/k_{\mathrm{B}}T_{\mathrm{S}})$ is the surface diffusion coefficient, D_{S0} is a constant, E_{S} ($=0.3$ eV) is the energy barrier for diffusion of carbon species on the catalyst surface [37], $D_{\mathrm{b}} = D_{\mathrm{b0}}\exp(-E_{\mathrm{b}}/k_{\mathrm{B}}T_{\mathrm{S}})$ is the bulk diffusion coefficient, D_{b0} is a constant, E_{b} ($=1.6$ eV) is the energy barrier for bulk diffusion of carbon species through catalyst nanoparticles [37], $D_{\mathrm{m}} = D_{\mathrm{m0}}\exp(-E_{SD}/k_{\mathrm{B}}T_{\mathrm{S}})$ is the diffusion coefficient of metal atoms [38] D_{m0} is a constant, E_{SD} is the activation energy of self-diffusion of metal atoms, p ($\approx$20 GPa) is the pressure exerted by graphene layers on the catalyst nanoparticles [16], $k = A_k\exp(-\delta E_{\mathrm{inc}}/k_{\mathrm{B}}T_{\mathrm{S}})$ is the incorporation speed of carbon species into the graphene layers [39], A_k is a constant, δE_{inc} is the energy barrier for diffusion of carbon species along the CNF–catalyst interface, ρ_{cnf} is the density of the CNF, $E_{a\mathrm{H}}$ ($=1.8$ eV for hydrogen) is the desorption energy, m_{cat} is the mass of the metal catalyst particle, and ρ_{cat} is the density of the metal catalyst particle.

Equation (23) accounts for the various processes to elucidate the growth mechanism of CNF on the metal catalyst surface. The first and second term in first bracket of the right-hand side represent the surface and bulk diffusion of carbon species through the catalyst surface, respectively. These diffusion processes significantly depend on the catalyst nanoparticle temperature. For lower catalyst temperatures, the surface diffusion process dominates the CNF growth whereas at higher catalyst temperatures the CNF growth is mainly due to the bulk diffusion process, and at intermediate temperature both the processes contribute to the CNF growth [40, 41]. However, in the present paper, we have considered both surface and bulk diffusion processes in CNF growth since the temperature of the catalyst nanoparticles is assumed to be constant. The third term represents the precipitation or incorporation of C species around the catalyst nanoparticles, i.e. realization of graphene layers around the catalyst nanoparticles. These carbon layers exert substantial stress on the catalyst nanoparticles, which can cause metal nanoparticles to deform. The catalyst nanoparticles no longer exhibited -spherical shape and appeared more like either a truncated cone or a teardrop. Now the question arises 'what stops the in plane growth of carbon condensation?', i.e. why carbon does not spread out over the whole substrate surface. The highly active atomic hydrogen and hydrogen ions etch the terminal carbons present at the edges of the graphitic nest and the carbon shadowed by the catalyst particle may contribute to the formation of graphitic shells around the catalyst particles or CNF. Therefore, the equation (23) also incorporates the adsorption of a hydrogen atom on the catalyst surface [37] (fourth term), the incoming flux of hydrogen ion due to thermal dehydrogenation of hydrocarbon ions [26] (fifth term), the ion induced decomposition of hydrogen [42] (sixth term), the flux of H-atoms due to thermal dissociation of hydrocarbons [42] (seventh term), direct flux due to hydrogen ion decomposition [47] (eighth term), and desorption of H-atoms from the catalyst surface (ninth term) [37]. The last term corresponds to metal diffusion [22], which is also responsible for the change in the catalyst particle morphology. Due to this morphological change and continuous carbon supply, the interaction between metallic catalyst nanoparticles and substrate decreases and hence the catalyst particles lift off, leading to the vertical growth of carbon nanofiber.

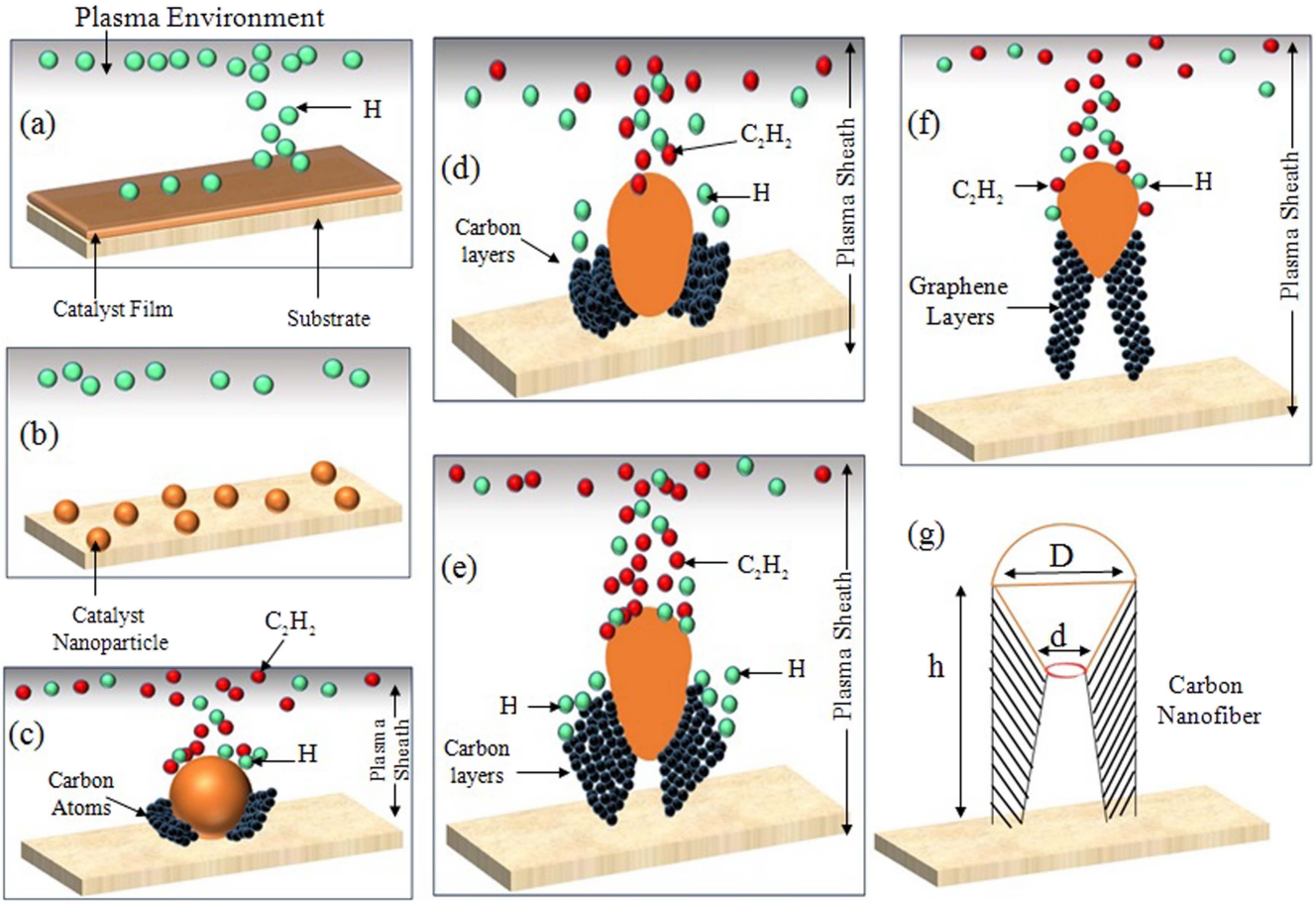


Figure 1. Schematic representation of the nucleation and growth mechanisms of CNF under plasma; (a) deposition of a thin catalyst film on the substrate, (b) formation of catalyst nanoparticles, (c) dissociation and adsoprtion of hydrocarbons and hydrogen on the catalyst particle surface, (d) precipitation of carbons and shape change of the catalyst particles, (e) detachment of catalyst particles from the substrate, (f) growth of CNF, and (g) scheme of the CNF considered.

3. Results and discussion

Figure 1 shows a schematic representation of different stages of nucleation and growth of carbon nanofiber on the catalyst–substrate surface in the presence of plasma. The growth pattern considered in figure 1 is similar to the growth of single- and multi-walled CNTs. Figure 1(a) shows the initial stage in which a thin catalyst film is deposited on the silicon substrate and kept in the plasma environment. We assume that the catalyst nanoparticles with spherical morphology are formed due to plasma etching (figure 1(b)). In figure 1(c), the dissociation and adsorption of the carbon gas precursor (C_2H_2) as well as hydrogen on the catalyst surface is shown. The carbon species diffuse through the bulk and surface of the catalyst particles and ultimately precipitate in the form of graphene layers at the rear surfaces of the catalyst nanoparticles (figure 1(c)). These graphene layers exert significant pressure on the catalyst particles, resulting in a shape change of the catalyst particles (figure 1(d)). The interaction (bonding force) between the catalyst nanoparticles and substrate is not very strong [43] and thus the catalyst nanoparticles are pushed up away from the substrate as the graphene layers precipitate out of the catalyst particles. Then, the active hydrogen and hydrogen ions etch the terminal carbons present at the edges of the graphitic nest and the carbon, shadowed by the catalyst nanoparticles, leading to the formation of new graphene layers (figure 1(e)). Due to the plasma sheath effect, the vertical growth continues [4, 26] and the catalyst particles change their shape continuously and finally transform to a pear-drop-like shape (figure 1(f)). The geometrical structure of the resulting CNF is shown in figure 1(g), where h is the height of the CNF, D is the initial diameter of the catalyst nanoparticles and d is the tip diameter or the inner diameter of the CNF.

Using MATHEMATICA software, the first-order differential equations (1)–(14) were solved simultaneously with appropriate boundary conditions viz., at $\tau = 0$, the electron number density n_{e0} ($=10^{12}\,\mathrm{cm}^{-3}$), ion number density n_{i10} ($=0.4n_{e0}$), n_{i20} ($=0.8n_{e0}$), neutral atom number density n_{10} ($=10^{13}\,\mathrm{cm}^{-3}$), n_{20} ($=10^{15}\,\mathrm{cm}^{-3}$), electron temperature T_{e0} ($=2\,\mathrm{eV}$), ion temperature T_{i0} ($=2500\,\mathrm{K}$), neutral temperature T_{n0} ($=2000\,\mathrm{K}$), mass of ion 1 m_{i1} ($=26$ amu for $C_2H_2^+$), mass of ion 2 m_{i2} ($=1$ amu for H^+), mass of neutral atom 1 m_i ($=26$ amu for C_2H_2), mass of neutral atom 2 m_2 ($=1$ amu for H), atomic mass of nickel $= 58.69$ amu, area of the catalyst film A_f ($=1\,\mathrm{cm}^2$), thickness of catalyst film d_f ($=10\,\mathrm{nm}$), density of catalyst Ni ρ_{ct} ($=8.96\,\mathrm{g\,cm}^{-3}$), coefficient of recombination of electrons and ions

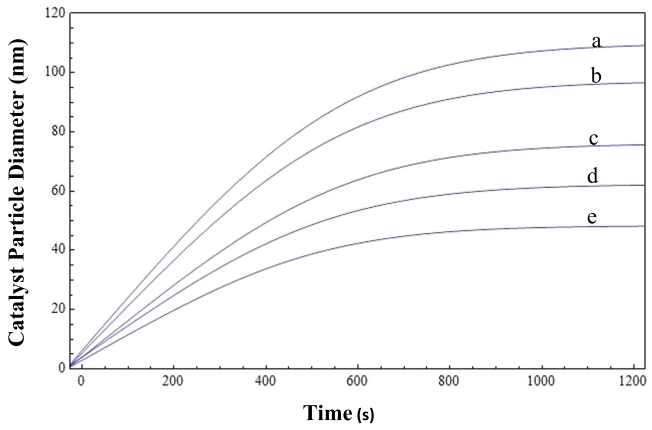


Figure 2. The variation with time of the catalyst particle diameter for different hydrogen ion number densities during the plasma pre-etching of thin film (where a, b, c, d and e correspond to $n_{i20} = 10^{11},\ 10^{12},\ 10^{13}\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3}$, respectively).

$\alpha_{10} \approx \alpha_{20}$ ($=1.12 \times 10^{-7}\,\mathrm{cm}^3\,\mathrm{s}^{-1}$), $\kappa = -1.2$, substrate temperature T_s ($=400\,^{\circ}\mathrm{C}$), sticking coefficient of an electron $\gamma_e(=1)$, sticking coefficient of ions $\gamma_{ij}(=1)$, total surface coverage $\theta_t = 0.01$, and temperature of the plasma environment $T_{\mathrm{env}} = 500\,^{\circ}\mathrm{C}$.

The etching gas density n_{20} is assumed to be greater than the hydrocarbon gas density n_{10} in order to avoid poisoning of the catalyst nanoparticles, i.e. formation of a thin layer of amorphous carbon on the front surface (the surface that is exposed towards the bulk plasma) of the catalyst nanoparticles. The poisoning of catalyst nanoparticles blocks the active surface for the decomposition of incoming hydrocarbons, and stops carbon nanofiber growth due to the loss of catalytic activity [44, 45].

Figure 2 illustrates the time evolution of catalyst nanoparticle size (diameter) obtained from the thin catalyst film for different hydrogen ion number density (i.e. $n_{i20} = 10^{11},\ 10^{12},\ 10^{13},\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3}$). From figure 2, it can be seen that the size of the catalyst nanoparticles increases with time and attains saturation value after some time. This figure also indicates the decrease in catalyst nanoparticle size with increase in hydrogen ion number density. This is attributable to the fact that, on increasing the hydrogen ion number density more etchant species are created, so that the etching of the film is increased as a result of heavy ion bombardment on the catalyst film surface, and, consequently, results in smaller catalyst nanoparticle size. Srivastava *et al* [20] claimed that, on increasing the plasma power, the density of the energetic plasma species increases. In this paper, we devised a model to relate the density of plasma species with input plasma power during plasma processing of thin catalyst film. Thus, the theoretical findings of figure 2 are in compliance with the experimental observations of Chang *et al* [19], Abdi *et al* [46] and Srivastava *et al* [20].

The variation of CNF length to the growth time for different catalyst nanoparticle size is shown in figure 3(i). It can be seen from figure 3(i) that as the catalyst nanoparticle size is increased, CNF of shorter length is observed. This signifies that one can get thick and shorter CNFs with large catalyst

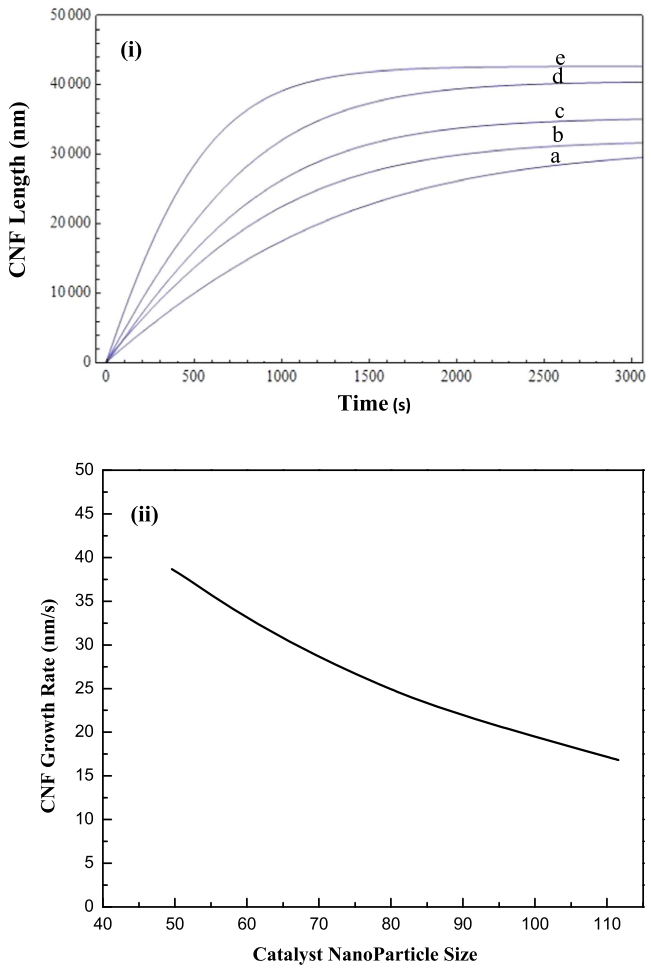


Figure 3. (i) The time evolution of CNF length for different catalyst nanoparticle sizes (where a, b, c, d and e correspond to $D = 109.3$, 96.34, 75.91, 61.95, 47.31 nm, respectively). (ii) The variation of the growth rate of CNF with catalyst nanoparticle size.

nanoparticles. The dependence of CNF length on the catalyst nanoparticle size during growth can also be interpreted in terms of the growth rate (CNF length per unit deposition time) of CNF, as shown in figure 3(ii). The growth rate of the CNF is decreased as the catalyst nanoparticle size is increased. This happens because for large catalyst nanoparticle size the carbon species diffusion length increases and therefore the growth of the CNF is slowed down. These results are in good agreement with the experimental observations given by Jong *et al* [47].

Figure 4 shows the variation of CNF length for different hydrocarbon number density ($n_{10} = 10^{11},\ 10^{12},\ 10^{13},\ 10^{14}\,\mathrm{cm}^{-3}$) in the plasma. Figure 4 indicates that the greater the hydrocarbon (C_2H_2) number density in the plasma the - larger the CNF length. This is explainable by the fact that with increasing number density of hydrocarbons, the dissociation of hydrocarbons at the catalyst nanoparticle surface increases resulting in greater diffusion of carbon species through the catalyst, which consequently increases the CNF length. The results of figure 4 are in accordance with the experimental observations of Matthews *et al* [14].

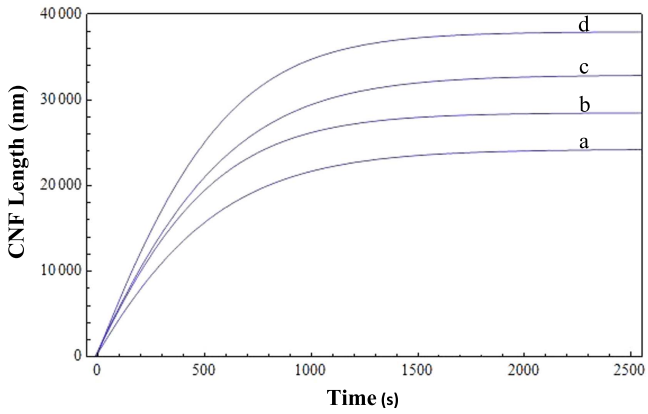


Figure 4. The time variation of the CNF length for different hydrocarbon number densities (where a, b, c, and d correspond to $n_{10} = 10^{11},\ 10^{12},\ 10^{13},\ 10^{14}\,\mathrm{cm}^{-3}$, respectively) at $n_{20} = 10^{15}\,\mathrm{cm}^{-3}$.

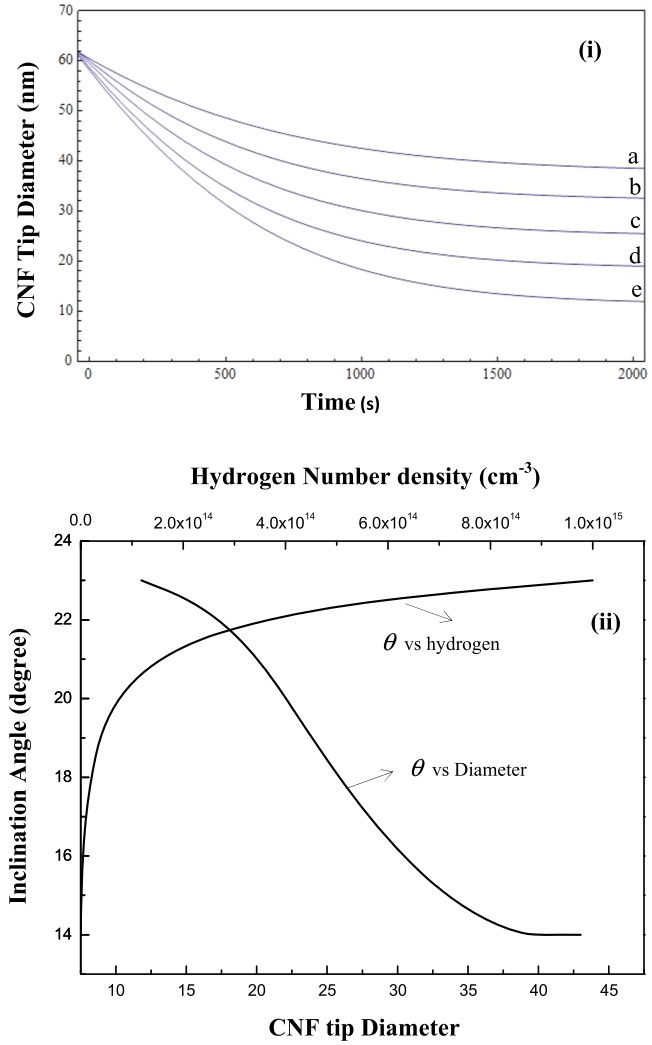


Figure 5. (i) The time evolution of CNF tip diameter for different hydrogen number density in the growth atmosphere (where a, b, c, d and e correspond to $n_{20} = 10^{11},\ 10^{12},\ 10^{13},\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3}$, respectively). (ii) The time variation of inclination angle with CNF tip diameter and hydrogen density.

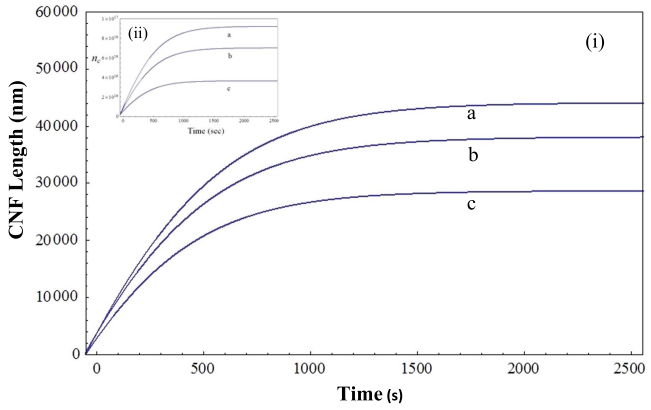


Figure 6. The time variation of the CNF length for different hydrogen number density (where a, b, and c correspond to $n_{20} = 10^{13},\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3}$, respectively) at $n_{10} = 10^{13}\,\mathrm{cm}^{-3}$.

Figure 5(i) shows the time evolution of the CNF tip diameter (CNF inner diameter) for different hydrogen number densities (i.e. $n_{20} = 10^{11},\ 10^{12},\ 10^{13},\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3}$). From figure 5(i), it can be seen that as the deposition time is increased, the inner diameter of the CNF is decreased. This happens because the growing graphene layers exert significant stress on the catalyst nanoparticles, which results in a change in the catalyst nanoparticle shape and consequently a decrease in the diameter of the base of the catalyst nanoparticles is observed. This is in good agreement with the experimental observations of Poa *et al* [48]. Furthermore, in figure 5(i), it is shown that an increase in hydrogen content $n_{20}(=10^{11},\ 10^{12},\ 10^{13},\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3})$ in the growth atmosphere decreases the inner diameter of the CNF. Two possibilities exist in this context: (i) more hydrogen adsorption over the metal catalyst particle surface weakens the metal–metal bond [23], which leads to the induced diffusion of metal atoms and consequently the base diameter of the catalyst nanoparticle or the inner diameter of the CNF tip is decreased, (ii) the abundant supply of hydrogen can terminate the large number of dangling bonds at the edges of the graphene layers and thus prevents the condensation reaction that leads to the formation of new graphene layers. Then, -catalyst nanoparticles experience more stress due to these layers and this leads to a decrease in catalyst particle size. As a result, one can think that the catalyst nanoparticle size and the relative inclination of graphene layers must be correlated (see appendix). Figure 5(ii) shows the dependence of the angle between the graphene layers and fiber axis on the catalyst nanoparticle base size as well as on the amount of hydrogen in the growth environment. It can be inferred from figure 5(ii) that an increase and decrease in hydrogen concentration and catalyst nanoparticle size, respectively, would tend to tilt the graphene layers away from the growth axis or fiber axis. The above observations are in accordance with the works of Nolan *et al* [49], Murayama and Maeda [50] and Cui *et al* [21].

The role of atomic hydrogen on CNF growth was also investigated in the present paper. In figure 6(i), the time variation of CNF length for different hydrogen number

densities ($n_{20} = 10^{13},\ 10^{14},\ 10^{15}\,\mathrm{cm}^{-3}$) is shown. As can be seen from figure 6(i), the CNF length decreases with increasing hydrogen atom number density in the plasma. This may be ascribed to the fact that, with an increase in the hydrogen atom number density, interaction between hydrogen and hydrocarbon on the catalyst nanoparticle surface (fifth term of equation (22)) increases, resulting in a decrease in the number (per unit area) of carbon species n_c generated on the catalyst nanoparticle surface, as shown in figure 6(ii). This leads to loss of carbon diffusion on the catalyst nanoparticle surface resulting in a decrease in CNF length and, eventually, a decrease in CNF growth rate. A similar trend was observed by other researchers [26, 31].

4. Conclusion

We introduced the analytical model for catalytic CNF growth in the PECVD process. The model accounted for the charging of CNFs, plasma sheath effects, kinetics of the electrons, ions and neutrals, and growth of CNF. Also taken into account were the effects relevant in catalytic CNF synthesis such as the formation of catalyst nanoparticles as a result of etching of thin catalyst film in the plasma atmosphere. It was shown that etching gas density, i.e. hydrogen number density, strongly affects the size of the catalyst nanoparticles. The results of our study show that the growth rate of CNF increases with a decrease in catalyst nanoparticle size. In addition, it was observed that upon increase in hydrocarbon number density, the CNF length increased. Moreover, we also showed the effect of hydrogen on the CNF tip diameter, CNF growth, and the inclination angle of graphene layers to the fiber axis. It was found that the tip diameter and CNF growth decreases and the graphene layers tilt away from the growth axis as hydrogen content in the growth atmosphere increases. This work could foresee the fabrication of CNF-based field emission devices in the near future.

Acknowldegement

One of the authors (RG) is thankful to Neha Gupta for her kind support and encouragement during the preparation of this manuscript. RG is grateful to the University Grant Commission (UGC), Government of India, for partial support of the work.

Appendix. Inclination angle of graphene layers to the growth axis

The geometry of the catalyst nanoparticle size is shown in figure 7, where l is the length of the truncated cone, D is the diameter of the catalyst nanoparticle and d is the diameter of the CNF tip. From figure 7 the angle between the graphene layers and the growth axis can be calculated in terms of the tip diameter

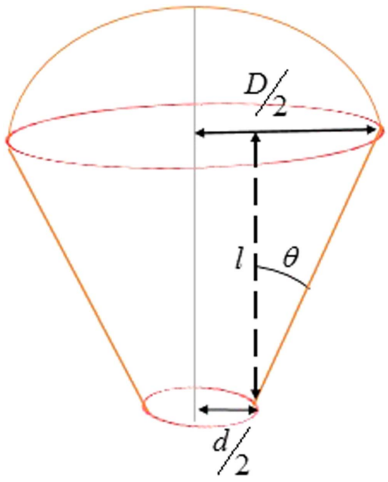


Figure 7. Geometrical representation of the deformed catalyst particles at the tip of the CNF.

of the CNF and the diameter of the particle size. Initially, the catalyst nanoparticles have spherical shape and finally transform to a pear-drop-like shape, i.e. hemispherical and truncated cone shape. Here, we are assuming that the volume of the catalyst nanoparticles remains constant during the CNF growth. Therefore, the initial volume must be equal to the final volume of catalyst nanoparticles.

Volume of sphere = volume of hemisphere + volume of truncated cone

$$\frac{1}{6}\pi D^3 = \frac{1}{12}\pi d^3 + \frac{1}{12}\pi (D^2 + D\,.\,d + d^2)l,$$

$$\text{and } \tan\theta = \frac{D-d}{2l},$$

$$\text{therefore, } \tan\theta = \frac{1}{2}\left[1 - \frac{d}{D} + \left(\frac{d}{D}\right)^2 - \left(\frac{d}{D}\right)^3\right] \text{ or}$$

$$\theta = \tan^{-1}\left(\frac{1}{2}\left[1 - \frac{d}{D} + \left(\frac{d}{D}\right)^2 - \left(\frac{d}{D}\right)^3\right]\right).$$

References

[1] Puretzky A A, Geohegan D B, Schittenhelm H, Fan X and Guillorn M A 2002 *Appl. Surf. Sci.* **197** 552
[2] Akita S, Kamo S and Nakayama Y 2002 *Jpn J. Appl. Phys. Part 2* **41** L487
[3] Cassell A M, Raymakers J A, Kong J and Dai H J 1999 *J. Phys. Chem.* B **103** 6484
[4] Merkulov V I, Melechko A V, Guillorn M A, Simpson M L, Lowndes D H, Whealton J H and Raridon R J 2002 *Appl. Phys. Lett.* **80** 4816
[5] Wei H W, Leou K C, Wei M T, Lin Y Y and Tsai C H 2005 *J. Appl. Phys.* **98** 044313
[6] Merkulov V I, Lowndes D H and Baylor L R 2001 *J. Appl. Phys.* **89** 1933
[7] Guillorn M A, Melechko A V, Merkulov V I, Ellis E D, Britton C L, Simpson M L, Lowndes D H and Baylor L R 2001 *Appl. Phys. Lett.* **79** 3506

[8] Murakami H, Hirakawa M, Tanaka C and Yamakawa H 2000 *Appl. Phys. Lett.* **76** 1776
[9] Dresselhaus M S, Dresselhaus G and Avouris P 2001 *Carbon Nano-tubes: Synthesis, Structure, Properties and Applications* (Berlin: Springer)
[10] Merkulov V I, Melechko A V, Guillorn M A, Lowndes D H and Simpson M L 2002 *Chem. Phys. Lett.* **361** 492
[11] Merkulov V I, Lowndes D H, Wei Y Y, Eres G and Voelkl E 2000 *Appl. Phys. Lett.* **76** 3555
[12] Ren Z F, Huang Z P, Wang D Z, Wen J G, Xu J W, Wang J H, Calvet L E, Chen J, Kelmic J F and Reed M A Reed 1999 *Appl. Phys. Lett.* **75** 1086
[13] Merkolov V I, Hensley D K, Melechko A V, Guillorn M A, Lowndes D H and Simpson M L 2002 *J. Phys. Chem.* B **106** 10570
[14] Matthews K, Cruden B A, Chen B, Meyyappan M and Delzeit L 2002 *J. Nanosci. Nanotechnol.* **2** 475
[15] Zhanbing H, Maurice J L, Gohier A, Lee C S, Pribat D and Cojocaru C S 2011 *Chem. Mater.* **23** 5379
[16] Yudasaka M, Kikuchi R, Matsui T, Ohki Y, Yoshimura S and Ota E 1995 *Appl. Phys. Lett.* **67** 2477
[17] Yang R T and Chen J P 1989 *J. Catal.* **115** 52
[18] Kim M S, Rodriguez N M and Baker R T K 1992 *J. Catal.* **134** 253
[19] Chang S C, Lin T C and Pai C Y 2007 *Microelectron. J.* **38** 657
[20] Srivastava S K, Shukla A K, Vankar V D and Kumar V 2005 *Thin Solid Films* **492** 124
[21] Cui H, Yang X, Simpson M L, Lowndes D H and Varela M 2004 *Appl. Phys. Lett.* **84** 4077
[22] Cojocaru C S, Senger A and Normand F L 2006 *J. Nanosci. Nanotechnol.* **6** 1
[23] Rodriguez N M 1993 *J. Mater. Res.* **8** 3233
[24] Zheng G B, Kouda K, Sano H, Uchiyama Y, Shi Y F and Quan H J 2004 *Carbon* **42** 635
[25] Tewari A and Sharma S C 2014 *Phys. Plasmas* **21** 063512
[26] Mehdipour H, Ostrikov K and Rider A E 2010 *Nanotechnology* **21** 455605
[27] Lieberman M A and Lichtenberg A J 1994 *Principles of Plasma Discharges and Material Processing* (New York: Wiley-Interscience)
[28] Ostrikov K N, Viladimirov S V, Yu M V and Morfill G E 2000 *Phys. Plasmas* **7** 461
[29] Viladimirov S V, Ostrikov K N and Yu M V 1998 *Phys. Rev.* E **58** 8046
[30] Keudell A V and Moller W 1994 *J. Appl. Phys.* **75** 7718
[31] Denysenko I and Ostrikov K 2009 *J. Phys. D: Appl. Phys.* **42** 015208
[32] Vattuone L, Yeo Y Y, Kose R and King D A 2000 *Surf. Sci.* **447** 1
[33] Kersten H, Deutsch H, Steffen H, Kroesen G M W and Hippler R 2001 *Vacuum* **63** 385
[34] Sodha M S and Guha S 1971 *Physics of Colloidal Plasmas* ed A Simon and W B Thomson (New York: John Wiley & Sons) p 219
[35] Sharma S C and Gupta N 2015 *Phys. Plasmas* **22** 123517
[36] Mantzaris N V and Gogolides E 1996 *J. Appl. Phys.* **79** 3718
[37] Denysenko I and Azarenkov N A 2011 *J. Phys. D: Appl. Phys.* **44** 174031
[38] See (http://matter.org.uk/matscicdrom/manual/df.html) for metal atom diffusion coefficient
[39] Louchev O A, Laude T, Sato Y and Kanda H 2003 *J. Chem. Phys.* **118** 7622
[40] Denysenko I and Ostrikov K 2007 *Appl. Phys. Lett.* **90** 251501
[41] Hofmann S, Csanyi G, Ferrari A C, Payne M C and Robertson J 2005 *PRL* **95** 036101
[42] Denysenko I, Ostrikov K, Yu M Y and Azarenkov N A 2007 *J. Appl. Phys.* **102** 074308
[43] Baker R T K, Barber M A, Waite R J, Harris P S and Feates F S 1972 *J. Catal.* **26** 51
[44] Hofmann S, Contoro M, Kleinsorge B, Casiraghi C, Parvez A, Robertson J and Ducati C 2005 *J. Appl. Phys.* **98** 034308
[45] Veríssimo C, Moshkalyov S A, Ramos A C S, Gonçalves J L, Alvesc O L and Swart J W 2006 *J. Braz. Chem. Soc.* **17** 1124
[46] Abdi Y, Arzi E and Mohajerzadeh S 2008 *Eur. Phys. J. Appl. Phys.* **44** 149
[47] Jong K P D and Geus J W 2000 *Catal. Rev. Sci. Eng.* **42** 481
[48] Poa C H P, Henley S J, Chen G Y, Adikaari A A D T, Giusca C E and Silvaa S R P 2005 *J. Appl. Phys.* **97** 114308
[49] Nolan P E, Lynch D C and Cutler A H Cutler 1998 *J. Phys. Chem.* B **102** 4165
[50] Murayama H and Maeda T 1990 *Nature* **345** 791

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2015): 78.96 | Impact Factor (2015): 6.391

Methods for the Screening and Treatment of Spinocerebellar Ataxias

Pushpendra Mani Mishra[1], Dr. Vimal Kishor Singh[2]

[1]Department of Biotechnology, Delhi Technological University, New Delhi 110042, India

[2]Founder & O/I, Stem Cell Research Laboratory, Department of Biotechnology, Delhi Technological University, New Delhi 110042, India

Abstract: ***Spinocerebellar ataxias (SCAs) are a heterogeneous group of inherited neurodegenerative disorders that occurs in autosomal dominant manner. The disease shares the similar molecular mechanism as other neurodegenerative disorders such as Alzheimer's and Parkinson's disease in which mutant protein form due to defective/ mutant DNA, these proteins accumulate inside neuronal cells and cause Neurodegeneration. The disorders affect mainly cerebellar regions, cerebellar connections, and its associated pathways in addition to other brain regions. The general characteristics of SCAs are cognitive impairment, depression, and motor deficits. The most prominent symptom observed in SCAs patients is ataxia with other varying non-ataxia symptoms in the majority of phenotypes. The modern classification strategies of the disorder is based upon expanding genotype. Due to overlap in clinical manifestation of many SCAs types, Genetic test, and various screening tools, for instance, Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Electromyography used to confirm the accumulation of mutant protein and aberrant neuronal mechanism responsible for the causation of particular phenotype. Nucleotides repeat expansion which is the basis of genetic disorder of SCAs can be determine by sequencing procedure. Similarly, techniques such as polymerase chain reaction (PCR) and gel electrophoresis with other molecular techniques and appliances used for the screening of patients in large number. The aim of this study was to explore the methods available for the screening of SCAs which form the basis of differential genetic diagnosis and list all the distinguishable clinical features and treatments available for SCAs.***

Keywords: Ataxia, Neuronal disorder, Spinocerebellar, Movement disorder, Neurodegenerative disease.

1. Introduction

The term ataxia refers to the loss of body balance and coordination while performing movements[1]. Conventionally, the autosomal dominant inherited ataxia is called spinocerebellar ataxia (SCA)[2]. Typical clinical features of SCAs are gait ataxia, appendicular ataxia followed by dysarthria, ocular difficulties in moving surrounding like saccadic eye movement, diplopia. Noncerebellar eye movement abnormalities are ocular "stare", blepharospasm, gaze palsies, ptosis and slowed saccades. Intention tremor on the finger to nose testing, dysmetria, widened stance and difficulty with gait, mainly when turning and rebound phenomena are characteristic cerebellar abnormalities[3].

The age of onset of disease is variable usually it is adult onset[4]. The worldwide prevalence of Autosomal dominant spinocerebellar ataxia is found to be 1to 5 cases per 100,000 population[5, 6]. Among all SCAs, SCA 3 is most common type which also known as Machado-Joseph disease (MJD) followed by SCA 1, 2, 6 and 7[7-21]. The site of mutation in different regions of the genome responsible for the numerous SCA types, several of the gene associated with this mutant site have already identified. Based upon the types of mutations, SCAs are categorized in 3 main group, group 1st is refers to SCAs with expanded CAG/polyQ ataxia, nonprotein coding repeat expansion ataxia comes under 2nd group and SCAs with conventional mutations (deletion, insertion, duplication, frameshift, and missense) kept in group 3rd. In SCAs most coalescing feature is a pattern of neurodegeneration, often clinician associate the disorder with clinical features which reflects the region of brain damage.[3] A growing body of evidence suggest the similar molecular mechanism behind the disease as other neurodegenerative diseases. Although based upon the types of mutation SCAs kept in 3 group but on the basis of genetic mechanism of disease, SCAs falls into four categories. Here it is noteworthy that the above 4 categories is for those SCAs in which the responsible gene has been identified already. The first category is SCAs with exonic triplet CAG repeats, for instance, SCA 1, 2, 3, 6, 7, 17, dentatorubropallidoluysian atrophy (DRPLA), and SCA 8 (partly), these codes for long, abnormal polyglutamine tracts within translated protein and results in a gain of toxic functions to these translated proteins. The second category is of SCAs with intronic or non-coding repeat expansion SCA 10, 12, 31, 36 and SCA 8(partly), some of these may involve in RNA-mediated toxicity. SCA 5, 11, 13, 14, 19/22, 23, 26, 27, 28, 29, and 35 with point mutations (including small insertions or deletions) may involve in a gain of novel function, haploinsufficiency or even dominant negative effect classified in category three of classification of the genetic mechanism. In last category the 4th one, SCAs with large duplications or deletions as SCA 15 and 20 have kept, they may result in gene dosage mechanisms i.e. under or over translation of the relevant protein products.[2]

2. Historical Perspective

The earliest practice of designating a disease with the term ataxia originate in the middle of the 19th century when Tabes dorsalis is named as Locomotor ataxia (a clinical term) by the French neurologist Duchenne de Boulogne. Autosomal dominant hereditary ataxia is referred as SCA, SCAs have assigned a different number based on its order of discovery. The term spinocerebellar ataxia first used in 1950 was based on the model of Friedreich ataxia. In the year, 1993 discovery of unstable CAG trinucleotide repeat expansion in SCA 1 disorder opens the way for the identification of mutations behind other autosomal dominant inherited SCAs.

Licensed Under Creative Commons Attribution CC BY

The method used previously for the identification and analysis of disease was linkage analysis. Gene discovery is the most recent method of identification of disorder. In the 20th century classification and neuropathological approaches to ataxia was based on the pivotal work of Greenfield and Holmes. The classification categories were based on the neuropathological characters such as spinocerebellar degeneration, or cerebellar cortical atrophy, olivopontocerebellar atrophy. However, this classification was not proper and the general consensus acceptance never reached. A new classification done by the professor Anita Harding in the beginning of 80s of the 20th century change the situation. SCAs were classified into four different types by Harding based on the specific characteristics, the four types are given as follows: The characteristics associated with Type 1 were cerebellar ataxia with optic atrophy, dementia, ophthalmoplegia, extrapyramidal signs, and amyotrophy. Disease symptoms ophthalmoplegia and extrapyramidal signs with retinal degeneration were kept in Type 2. The third category of classification (Type 3) was implicated with pure cerebellar ataxia. Type 4 was defined with the symptoms myoclonia and deafness in addition to cerebellar ataxia. Figure-1 exhibit most important events of ataxia on the historical timeline. Table-1 articulate the description of these important major historical proceedings on ataxia.[22, 23]

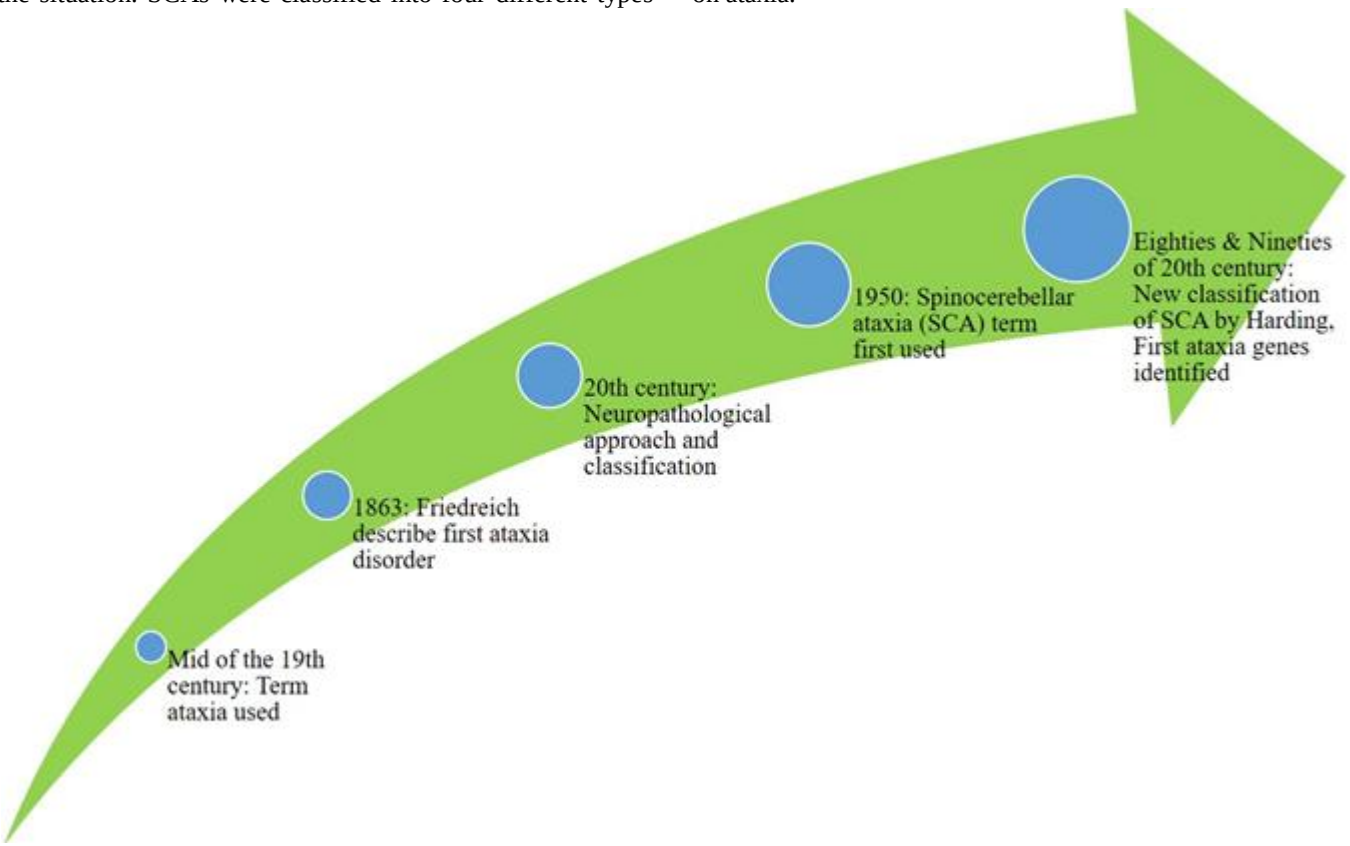


Figure 1: Most important events of ataxia on historical timeline

Table 1: Most important Historical occurrences of ataxia

Period	***Historical occurrence***
Mid of the 19th century	Term ataxia used to designate a disease (Tabes dorsalis termed as locomotor ataxia)
1863	Friedreich describe in details the first ataxia disorder
20th century	Neuropathological approach to ataxia and Neuropathological category based classification of ataxia established
1950	A historical term first used in 1950s based on Friedreich ataxia as a model
Eighties of 20th century	New classification of SCA by Harding based on genetic and clinical criteria
1993	Identification of mutation causing dominantly inherited SCAs started, unstable expansion of a translated CAG repeats cause SCA-1, identified
Nineties of 20th century	Clinical research establish that the SCA-1, SCA-2, SCA-3 Patients demonstrate additional symptoms in addition to ataxia while SCA-6 is an almost pure cerebellar ataxia

3. Symptoms and Visible Characteristics

SCAs are a group of complex disease which is heterogeneous in genotypic as well as phenotypic sense. The phenotypic heterogeneity implies that the different phenotype is determined by the same genotype, genotypic heterogeneity defined as the presence of the same phenotype although the genotypes are different.[24] The modern or new classification of ataxia is based on the order of discovery of locus of mutation. The number of genes have determined which implicated with these discovered locus although few of them yet to be determined, the most common features related to disease are uncoordinated movement, eye and speech impairments, cognitive decline, neuropathy, and spasticity.[4, 25] The current status of genetic testing is limited and cannot offer testing for all types of ataxia. The initial decision of genetic testing for a particular type of ataxia based on the clinical features that differentiate among many types, patient disease history and family disease history.[26] Table-2 summarizes the particular gene/locus, mutation, average onset time, average disease duration and clinical

Licensed Under Creative Commons Attribution CC BY

features associated with SCAs types. The clinical symptoms is also included non-ataxia symptoms in addition to the ataxia symptoms associated with different SCAs.

Table 2: SCA types, their genes, mutations, average onset, average duration and distinguished clinical features

SCAs Types	*Gene/ Locus*	*Mutation*	*Ranges in Years (Average onset)*	*Ranges in Years (Average durations)*	*Distinguish Clinical Features*
SCA 1	*ATXN1*	CAG repeat	<10 to>60 (3rd-4th decade)	10-28 (15)	**Pyramidal signs**, Peripheral neuropathy,
SCA 2	*ATXN2*	CAG repeat	<10 to>60 (3rd-4th decade)	1-30 (10)	**Slow saccades**, Peripheral neuropathy, Decreased DTRs, Dementia
SCA 3	*ATXN3*	CAG repeat	10-70 (4th decade)	1-20 (10)	**Pyramidal sign**, Amyotrophy fasciculations, sensory loss, Lid retraction, nystagmus, and decreased saccade velocity,
SCA 4	16q22.1	Unidentified	19-72 (4th-7th decade)	(Decades)	**Sensory axonal neuropathy,** Deafness
SCA 5	*SPTBN2*	Missense, in frame deletion	10-68 (3rd-4th decade)	(>25 years)	**Pure cerebellar ataxia (late onset)**, Pyramidal sign (early onset)
SCA 6	*CACNA1A*	CAG repeat	19-71 (5th-6th decade)	(>25 years)	**Pure cerebellar ataxia late onset**
SCA 7	*ATXN7*	CAG repeat	0.5-60 (3rd-4th decade)	Early onset correlate with shorter duration, 1-45 (20)	**Pigmentary macular degeneration**
SCA 8	*ATXN8*	CTG & CAG repeat	1-65 (4th decade)	(Normal life span)	**Pyramidal sign**, Sometimes brisk DTRs, decreased vibration sense, Rarely, cognitive impairment, Slowly progressive
SCA 9	*Not Assigned*				
SCA 10	*ATXN10*	ATTCT Penta nucleotide expansion	12-48 (4th decade)	(9 years)	**Seizures**
SCA 11	*TTBK2*	Frameshift	15-70 (age 30 years)	(Normal life span)	**Pure cerebellar ataxia**
SCA 12	*PPP2R2B*	CAG repeat	8-62 (4th decade)		**Upper extremity tremor**, Slowly progressive ataxia, Subtle parkinsonism possible, Action tremor in the 30s, Hyperreflexia, Cognitive/psychiatric disorders including dementia
SCA 13	*KCNC3*	Missense	Adulthood or childhood	Unknown	**Intellectual disability**, Short stature
SCA 14	*PRKCG*	Missense	3-70 (3rd-4th decade)	1-30 (Decades)	**Pure cerebellar ataxia,** Early axial myoclonus
SCA 15	*ITPR1*	Deletion & Missense	7-66 (4th decade)	(Decades)	**Pure cerebellar ataxia**
SCA 16	*ITPR1*	Deletion & Missense	20-66 (age 39 years)	(1-40 years)	**Pure cerebellar ataxia,** Head tremor
SCA 17	*TBP*	CAG repeat	3-55 (4th decade)	(>8 years)	**Chorea**, dystonia, myoclonus, epilepsy, Mental deterioration
SCA 18	7q22-q32	Unidentified	12-25 (adolescence)	(Decades)	**Posterior column loss**, Decreased tendon reflexes, Dysarthria, Nystagmus, Ataxia with early sensory/motor neuropathy
SCA 19/22	*KCND3*	Deletion & Missense	10-51 (4th decade)	(Decades)	**Pure cerebellar ataxia,** slowly progressive, Myoclonus, Hyperreflexia, rare cognitive impairment
SCA 20	11q12.2-11q12.3	Duplication	19-64 (5th decade)	(Decades)	**Spasmodic dysphonia** or **Spasmodic coughing**, Bradykinesia, Hyperreflexia, Early dysarthria
SCA 21	*TMEM240*	Missense &Nonsense	30-Jun	(Decades)	Mild to severe early-onset cognitive impairment
SCA 23	*PDYN*	Missense	(5th-6th)	(>10 years)	**Late onset**, Abnormal eye movements, Dysarthria, Reduced vibration and position sense
SCA 25	*SCA25*	Unidentified	1.5-39	Unknown	Sensory neuropathy
SCA 26	*EEF2*	Missense	26-60	Unknown	Irregular visual pursuits, Dysarthria
SCA 27	*FGF14*	Missense & Frameshift	7-20 (age 11 years)	(Decades)	Orofacial dyskinesia, Cognitive deficits, Early-onset tremor
SCA 28	*AFG3L2*	Missense	12-36 (age 19.5 years)	(Decades)	**Early onset**, Increased tendon reflexes, Ptosis, Ophthalmoparesis, Nystagmus
SCA 29	*ITPR1*	Missense	Early Childhood	(Lifelong)	**Congenital**, Learning deficits
SCA 30	4q34.3-q35.1	Unidentified	45-76	(Lifelong)	Hyperreflexia
SCA 31	*BEAN1*	TGGAA Penta nucleotide repeat	61.2 (3rd-4th decade)	Unknown	**Pure cerebellar ataxia** late onset, usually >50 years
SCA 34	*ELOVL4*	Missense	Early Childhood (4th or 5th decade)	Unknown	Skin changes disappear in adulthood
SCA 35	*TGM6*	Missense &deletion	Teenage to late	Unknown	**Pyramidal sign**

Licensed Under Creative Commons Attribution CC BY

			adulthood		
SCA 36	*NOP56*	GGCCTG Hexanucleotide Repeat	Adulthood	Unknown	**Lower motor neuron involvement**, Hyperreflexia, Tongue atrophy, Muscle fasiculations
SCA 37	1p32	Unidentified	Adulthood	Unknown	Abnormal vertical eye movements
SCA 38	*ELOVL5*	Missense	Adulthood	Unknown	Axonal neuropathy, Adult onset
SCA 40	*CCDC88C*	Missense	Adulthood	Unknown	Brisk reflexes, Spasticity, Adult onset
DRPLA	*ATN1*	CAG repeat	20	Unknown	**Myoclonic epilepsy**, Dementia, Seizures, Chorea
SCA 42	*CACNA1G*	Missense	Highly Variable	Unknown	Saccadic pursuit, Mild pyramidal signs

·DTR, deep tendon reflex
Symptoms in Bold typeface represent unique or characteristic feature of the particular SCA type

4. Methods for screening of SCAs

Numerous methods of identification of different types of SCAs applied in hospitals and research laboratory for screening, Figure-2 display the available methods for primary/ initial screening of different types of SCAs. Theses screening methods helpful in the construction of Algorithm for differential genetic diagnosis of SCAs.



Figure-2 Screening methods helpful in the construction of differential genetic testing algorithm for SCAs

The detail descriptions of screening methods of SCAs are as follows:

4.1 Electrophysiological

The basis of this method is an electrophysiological measurement of masseter reflex. The initial study was done on the patients of SCA type 2 and 3 by Antonio Garcia et al., the device used for the study was electromyography (EMG) system (Medelec-Synergy EMG/PE; Cardinal Health, UK). Masseter reflex measurement is a reliable test that discriminates between the SCA type 2 and 3, although this method was used for primary screening purpose.[27]

4.2 Radiological

A study carried by Yoshio Murata et al. found the certain characteristics feature of the brain of MJD patients in MRI study. The MRI study demonstrates the atrophy of temporal and frontal lobe in addition to Globus pallidus, abnormality in afferent and efferent cerebellar tracts also reported. These features may helpful in finding difference between sporadic olivopontocerebellar atrophy (sOPCA) and MJD.[28] Another published work of Yoshio Murata et al. signify the distinguishable feature of SCA 6 from other SCAs forms, in an MRI study of the brain. The MRI study shows the affected cerebellar region and its afferent and efferent system.[29] MRI is the steadfast method to find the affected brain regions in SCAs patient and other neurodegenerative disorder but it is not a confirmatory test for SCAs types, due to several overlap in clinical characteristics.

4.3 Biochemical

A study done by Marcelle R. Morrison and Roger N. Rosenberg on the brain of Machado-Joseph disease patients led the finding that there were an elevation in steady state level of many proteins such as glial fibrillary acidic protein complex, J protein complex, and L protein complex in the cerebellar and cortex region of brain in comparison to the control brain protein analysis.[30]

Licensed Under Creative Commons Attribution CC BY

4.4 Radiochemical

An comparative analysis of dopamine transporter and regional cerebral glucose metabolism (rCMRglu) by using [11C]d-threo-methylphenydate ([11C]dMP) and [18F]fluorodeoxyglucose ([18F]FDG) respectively and Positron emission tomography (PET) screening carried by Ullrich Wüllner et al., in patients with SCA 1, 2, 3 and 6 in comparison to Parkinson disease (PD) patients and healthy controls. Their findings reflect the specific pattern of neuronal susceptibility met in each polyglutamine disorder. The specific pathological features observed in patients with SCA 1, SCA 2, SCA 3 and SCA 6 provide a stepping stone for the identification of disorder among several pathotypes. They observed no significant change in the striatal [11C]dMP binding potential (BPdMP) but reduce rCMRglu in the cerebellum and brain stem of SCA 1 patients. . In SCA 2 patient's striatal BPdMP was evidently reduced, putamen and caudate nucleus region showed the 57 and 55% of BPdMP reduction respectively. An increment in the rCMRglu was noted in the temporal cortical region of interest (ROI) analysis. The dopamine terminal (DAT) loss was throughout striatum and similar to PD with a more severely affected putamen. Particularly in the case of SCA 2 mutation, the dopaminergic system was sensitive. Pathological features of SCA 3 demonstrated the marked reduction of striatal BPdMP. The same degree of reduction of BPdMP in the putamen (29%) and caudate (20%) were observed. Brainstem, Thalamus and putamen region of brain displayed, the reduce level of rCMRglu, the area of decreased metabolism extended from cerebellar midline structures to adjacent pons and midbrain. An increased metabolism was observed in the superior and middle temporal gyri. The loss of DAT observed in SCA 3 patients but it was less severe in comparison to patients with SCA 2. The loss of dopaminergic cells and global synaptic impairment might contribute to the alterations of DATs in SCA 3 patients. Reduced rCMRglu level was observed in the putamen region of patients with SCA 6. Differentiated brain regions with the increase and decrease metabolic activity were observed in temporal gyri (superior, middle) and Cerebellum respectively in patients with SCA 6.[31] A study done by Bing-wen Soong and Ren-shyan Liu in asymptomatic gene carriers of MJD demonstrate a regional decrease in brain glucose metabolism. They used positron emission tomography (PET) with [18F]-2-fluoro-2-deoxy-D-glucose (FDG) to record the metabolic changes in the asymptomatic carrier.[32]

4.5 Molecular and Genetic

Molecular and genetic testing techniques depend on the diverse types of mutation implicated in different SCAs types. The most common type of mutation is trinucleotide (CAG) repeats found in coding regions of genes in SCA 1, 2, 3, 6, 7, 17 and DRPLA. Bidirectional transcribed CTG and CAG repeats in 3'UTR ATXN8OS region as well as ORF ATXN8 region found in SCA 8. The nontranslated mutation found in SCA 10, SCA 12 and SCA 31, pentanucleotide repeat ATTCT and TGGAA found in the intronic region of SCA 10 and SCA31 respectively, CAG repeats present in 5'UTR of SCA 12. Hexanucleotide repeats GGCCTG found in the intronic region of genes in SCA 36. The conventional mutations (missense, deletion, duplication and frameshift) has reported in SCA type 5, 11, 13, 14, 15, 16, 20,21,23,26, 27, 28, 29,34, 35, 38, 40 and 42 and DRPLA.[23, 33-50] Table-2 demonstrate the particular mutation and associated SCAs types. The most reliable and popular method of screening of spinocerebellar ataxia is Genetic diagnosis which relies upon Genetic material DNA and molecular techniques. The molecular genetic screening method used in different parts of the world in various races of Human population. In most of the SCAs where the mutation is due to nucleotides repeats, PCR screening technique is used to amplify the mutated region of the genomic fragment, analysis of amplified region is done with gel electrophoresis or automated sequencer.[51-56]

The other molecular techniques include repeat expansion detection (RED) method, Linkage analysis, Whole-Exome sequencing, Whole genome sequencing and whole mitochondrial sequencing. Kerstin Lindblad et al. used the RED method to detect CAG repeats in SCA 7 patients and in their family, in the RED method they used thermostable DNA ligase to detect repeats directly from DNA.[57] Lourdes Martorell et al. detected CAG/CTG expansion in SCA 1 and MJD patients by using the RED method.[58] The detection of CAG repeat expansion in MJD patients by RED method demonstrated in another publication by Kerstin Lindblad et al.[59] A published work of Kokoro Ozaki et al. and Je´roˆme Delplanque et al. demonstrate the use of linkage analysis in determination of mutation in SCA 34 and SCA 21 respectively. In the Genome-wide linkage analysis, they used SNP array chip for mutation detection. Most of the SCAs which cannot be identified by molecular techniques such as PCR, RED, and linkage analysis method can identify by Whole exome sequencing, whole genome sequencing or whole mitochondrial sequencing. Sequencing technique most preferably used for the identification of Conventional mutations linked SCAs.[23, 36, 37] Numerous combination of molecular techniques and appliances which also include variants of PCR are used for the genetic screening of SCAs.[60-72]

5. Algorithm for the differential genetic diagnosis of SCAs

Differential genetic diagnosis of SCAs depend upon the number of factors, to proceed for the genetic diagnosis it is very much important that the details of clinical history, family history and outcomes of all neurological examination must be known. The first step is to rule out all the acquired (non-genetic) cause of ataxia such as vitamin deficiencies, multiple sclerosis, vascular disease, primary or metastatic tumors, alcoholism, or paraneoplastic diseases associated with occult carcinoma of the breast, ovary, or lung and then proceed for the genetic testing. Figure-3 describe the steps involve in differential genetic diagnosis of SCAs.[23, 73, 74]

Licensed Under Creative Commons Attribution CC BY

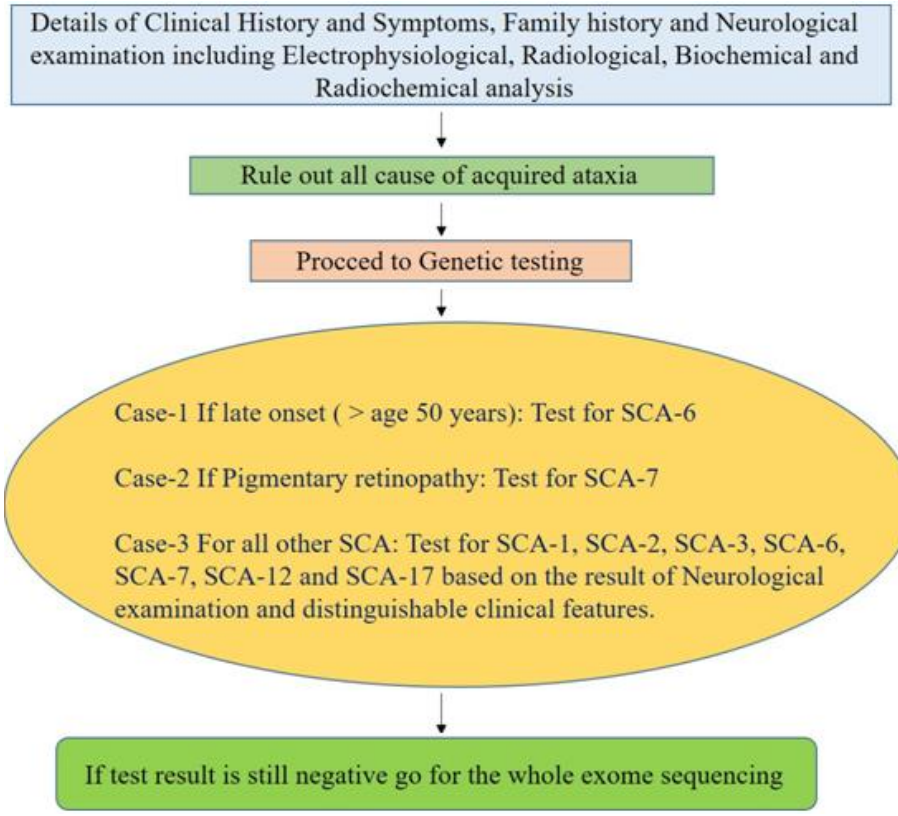


Figure 3: Algorithm for differential genetic diagnosis of SCAs

6. Approaches for the treatment of SCAs

In most of hereditary ataxia, only supportive treatment is possible these includes centrally acting drugs such as 5-hydroxytryptophan, buspirone and physostigmine and thyrotropin-releasing hormone and D-cycloserine.[22] To develop the definite and possible cure for ataxia patients many approach and different experimentation have been carried out and ongoing. A study done by Song Tan et al., in SCA 3 patients determined the potential of Nerve growth factor for the treatment of SCA 3. A clinical assessment of effect of N-acetylcysteine on 18 patients of Spinocerebellar disease by Dr. Roswell Eldridge demonstrate a marked improvement in degrees of ataxia, dysarthria and oculomotor disturbance from disease condition. Jia-Li Jin et al., experimentation on 16 SCAs patients (Genetically diagnosed and confirmed) including SCA-1, SCA 2 and SCA 3 establish the improvement in motor and balancing ability after receiving intrathecal and intravenous infusion of umbilical cord Mesenchymal Stem cells (MSCs). The laboratory examinations demonstrated that the umbilical cord-MSC therapy was safe. In addition to clinical study of human subjects, In vivo and In vitro studies on Cell lines and Animal Models determine the possibilities of cure in Human subjects. A study by Vladimir F. Lazarev et al., in Cell and Fly model of spinocerebellar ataxia 3 prove the Neuroprotective role of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) binding drugs such as Deprenyl (DEP) and 2-amino-4,7 dimethyl benzothiazole (PGL). GAPDH form the complex with protein with long polyglutamine repeats. Additional approaches which could be possible for the treatment of ataxia enlisted by David D. Bushart et al., in his review paper, These are (1) RNA interference and Antisense Oligonucleotide-based strategies (2) Heat Shock Protein (Hsp) 70 and Hsp 27 activation (3) Use of Histone deacetylase (HDAC) Inhibitor (4) Expression of Calcium binding protein and Inhibitor of Intracellular Calcium release store (5) Dantrolene and Glutamate transport activator (6) Potassium channel modulator and Calcium-activated potassium channel activators. The RNA interference or Antisense oligonucleotide strategies have been used in mouse model of SCA 1, SCA 3 and SCA 7 for silencing of specific mutant allele which leads to the improvement in condition from disease state. It may have relevance for the treatment of SCAs in human subjects. Lymphoblastoid cells derived from SCA 7 patient show reduced expression of Hsp70 and Hsp27, activation of these two proteins may have neuroprotective benefit. In SCA 3 mouse model H3 and H4 histone found to be hypo acetylated it suggest the presence of HDAC over activity, use of HDAC inhibitor Sodium butyrate in SCA 3 mouse model improve Histone acetylation and increase the expression of gene which were suppressed earlier in cerebellum region. HDAC inhibitors may have therapeutic promise for the treatment of some forms of SCA, particularly those which display repressed gene transcription condition. Drug-mediated treatments which increase the glutamate uptake, increase expression of calcium binding proteins, or inhibit intracellular calcium release store and modulate potassium channel might be promising candidates for treating neuronal dysfunction and synaptic physiology across different cause of SCAs. Table-3 summarize the possible available and continuing treatment methods for SCAs with brief description.[75-81]

Licensed Under Creative Commons Attribution CC BY

Table 3: Possible treatment types of SCA and their descriptions

Treatment Types	Description
Deprenyl (DEP) and 2-amino-4,7-dimethylbenzothiazole (PGL)	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) binding drugs have neuro-protective activity in cell and fly (Drosophila) model of SCA-3
Human Umbilical derived Mesenchymal Stem cells (MSCs)	Intrathecal and Intravenous injection of Umbilical MSCs to Human ataxia patients (SCA-1, SCA-2 and SCA-3) improve the motor and balancing ability.
N-acetylcysteine	Medication in Spinocerebellar disease patients demonstrate improvement in degrees of ataxia, dysarthria, and oculomotor disturbance from disease condition.
Nerve growth factor (NGF)	Intramuscular injection effective in treating patients with SCA3.
RNA interference based and Antisense oligonucleotide based strategies	Silence specific alleles may have relevance for the treatment of polyglutamine in SCAs.
Heat shock protein (Hsp) 70 and Hsp 27 activation	Hsp70 and Hsp27activators may promote clearance and degradation of polyglutamine proteins, have neuroprotective benefits in SCAs
Histone deacetylase (HDAC) inhibitor	HDAC inhibitors may have therapeutic promise for SCA which display repressed gene transcription.
Molecules increase expression of calcium binding protein or inhibit intracellular calcium release store	It may improve neuronal health and limit calcium-mediated neuronal toxicity
Dantrolene and glutamate transport activator	Improve synaptic physiology
Small-conductance calcium-activated potassium (SK) channel activators	Treatment of neuronal dysfunction

7. Conclusion

This review illustrates the clinical manifestation and available diagnostic methods and treatment of neurodegenerative disorder SCAs. It is very necessary for the individual patients and their family to come to an appropriate diagnosis to confirm the disorder and affected regions of the body. The gold standard test for SCAs is a Molecular and Genetic diagnosis, Although Physiological, Radiological, and Radiochemical methods are not confirmatory tests but it is helpful in the determination of affected brain region and extent of damage occurs to the brain and construction of algorithm for the differential genetic diagnosis of SCAs. These test may also confirm the anticipated consequence to the body as the Nervous system is a master controller of organs and organs associated functions. This review also summarize the outcome of various experimental approaches for the treatment of SCAs.

References

[1] Orr HT. “Cell biology of spinocerebellar ataxia,” J Cell Biol., 197(2), pp. 167-177, 2012.
[2] Storey E. “Genetic cerebellar ataxias,” Semin Neurol., 34(3), pp. 280-292, 2014.
[3] Paulson HL. “The spinocerebellar ataxias,” J Neuroophthalmol., 29(3), pp. 227-237, 2009.
[4] Jayadev S, Bird TD. “Hereditary ataxias: overview,” Genet Med., 15(9), pp. 673-683, 2013.
[5] van de Warrenburg BP, Sinke RJ, Verschuuren-Bemelmans CC, Scheffer H, Brunt ER, Ippel PF, Maat-Kievit JA, Dooijes D, Notermans NC, Lindhout D, Knoers NV, Kremer HP. “Spinocerebellar ataxias in the Netherlands: prevalence and age at onset variance analysis,” Neurology., 58, pp. 702–708, 2002.
[6] Ruano L, Melo C, Silva MC, Coutinho P. “The global epidemiology of hereditary ataxia and spastic paraplegia: a systematic review of prevalence studies,” Neuroepidemiology., 42, pp. 174–183, 2014.
[7] Schöls L, Amoiridis G, Buttner T, Przuntek H, Epplen JT, Riess O. “Autosomal dominant cerebellar ataxia: phenotypic differences in genetically defined subtypes?,” Ann Neurol., 42, pp. 924–932, 1997.
[8] Moseley ML, Benzow KA, Schut LJ, Bird TD, Gomez CM, Barkhaus PE, Blindauer KA, Labuda M, Pandolfo M, Koob MD, Ranum LP. “Incidence of dominant spinocerebellar and Friedreich triplet repeats among 361 ataxia families,” Neurology., 51, pp. 1666–1671, 1998.
[9] Saleem Q, Choudhry S, Mukerji M, Bashyam L, Padma MV, Chakravarthy A, Maheshwari MC, Jain S, Brahmachari SK. “Molecular analysis of autosomal dominant hereditary ataxias in the Indian population: high frequency of SCA2 and evidence for a common founder mutation,” Hum Genet., 106, pp. 179–187, 2000.
[10] Storey E, du Sart D, Shaw JH, Lorentzos P, Kelly L, McKinley Gardner RJ, Forrest SM, Biros I, Nicholson GA. “Frequency of spinocerebellar ataxia types 1, 2, 3, 6, and 7 in Australian patients with spinocerebellar ataxia,”. Am J Med Genet., 95, pp. 351–357, 2000.
[11] Tang B, Liu C, Shen L, Dai H, Pan Q, Jing L, Ouyang S, Xia J. “Frequency of SCA1, SCA2, SCA3/MJD, SCA6, SCA7, and DRPLA CAG trinucleotide repeat expansion in patients with hereditary spinocerebellar ataxia from Chinese kindreds,” Arch Neurol., 57, pp. 540–544, 2000.
[12] Maruyama H, Izumi Y, Morino H, Oda M, Toji H, Nakamura S, Kawakami H. “Difference in disease-free survival curve and regional distribution according to subtype of spinocerebellar ataxia: a study of 1,286 Japanese patients,” Am J Med Genet., 114, pp. 578–583, 2002.
[13] Silveira I, Miranda C, Guimaraes L, Moreira MC, Alonso I, Mendonca P, Ferro A, Pinto-Basto J, Coelho J, Ferreirinha F, Poirier J, Parreira E, Vale J, Januario C, Barbot C, Tuna A, Barros J, Koide R, Tsuji S, Holmes SE, Margolis RL, Jardim L, Pandolfo M, Coutinho P, Sequeiros J. “Trinucleotide repeats in 202 families with ataxia: a small expanded (CAG)n allele at the SCA17 locus,” Arch Neurol., 59, pp. 623–629, 2002.
[14] van de Warrenburg BP, Sinke RJ, Verschuuren-Bemelmans CC, Scheffer H, Brunt ER, Ippel PF, Maat-Kievit JA, Dooijes D, Notermans NC, Lindhout D, Knoers NV, Kremer HP. “Spinocerebellar ataxias in the Netherlands: prevalence and age at onset variance analysis,” Neurology., 58, pp. 702–708, 2002.

Licensed Under Creative Commons Attribution CC BY

[15] Dryer SE, Lhuillier L, Cameron JS, Martin-Caraballo M. "Expression of K(Ca) channels in identified populations of developing vertebrate neurons: role of neurotrophic factors and activity," J Physiol Paris., 97, pp. 49–58, 2003.
[16] Brusco A, Gellera C, Cagnoli C, Saluto A, Castucci A, Michielotto C, Fetoni V, Mariotti C, Migone N, Di Donato S, Taroni F. "Molecular genetics of hereditary spinocerebellar ataxia: mutation analysis of spinocerebellar ataxia genes and CAG/CTG repeat expansion detection in 225 Italian families,". Arch Neurol., 61, pp. 727–733, 2004.
[17] Schöls L, Bauer P, Schmidt T, Schulte T, Riess O. "Autosomal dominant cerebellar ataxias: clinical features, genetics, and pathogenesis," Lancet Neurol. 3, pp. 291–304, 2004.
[18] Shimizu Y, Yoshida K, Okano T, Ohara S, Hashimoto T, Fukushima Y, Ikeda S. "Regional features of autosomal-dominant cerebellar ataxia in Nagano: clinical and molecular genetic analysis of 86 families," J Hum Genet., 49, pp. 610–616, 2004.
[19] Zortea M, Armani M, Pastorello E, Nunez GF, Lombardi S, Tonello S, Rigoni MT, Zuliani L, Mostacciuolo ML, Gellera C, Di Donato S, Trevisan CP. "Prevalence of inherited ataxias in the province of Padua, Italy,". Neuroepidemiology., 23, pp. 275–280, 2004.
[20] Jiang H, Tang BS, Xu B, Zhao GH, Shen L, Tang JG, Li QH, Xia K. "Frequency analysis of autosomal dominant spinocerebellar ataxias in mainland Chinese patients and clinical and molecular characterization of spinocerebellar ataxia type 6," Chin Med J (Engl), 118, pp. 837–843, 2005.
[21] Jiang H, Wang J, Du, J, Duan R, Li J, Tang B. "Progress in treating hereditary ataxia in mainland China. In: Sanders S, Zhang Z, Tang V, eds. Pathways to Cures: Neurodegenerative Diseases in China,". Washington, DC: Science/AAAS, pp. 32-34, 2013.
[22] Klockgether T, Paulson H. "Milestones in ataxia. Mov Disorder," 26(6), pp. 1134-1141, 2011.
[23] Bird TD. Hereditary Ataxia Overview. 1998 Oct 28 [Updated 2016 Mar 3]. In: Pagon RA, Adam MP, Ardinger HH, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2016. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1138/
[24] Teive HA. "Spinocerebellar ataxias," Arq Neuropsiquiatr., 67(4), pp. 1133-1142, 2009.
[25] Trott A, Houenou LJ. "Mini-review: spinocerebellar ataxias: an update of SCA genes," Recent Pat DNA Gene Seq., 6(2), pp. 115-121, 2012.
[26] Tan EK, Ashizawa T. "Genetic testing in spinocerebellar ataxias: defining a clinical role,"Arch Neurol., 58(2), pp. 191-195, 2001.
[27] García A, Alvarez S, Infante J, Berciano J. "Masseter reflex in the study of spinocerebellar ataxia type 2 and type 3," Muscle Nerve., 40(4), pp. 640-642, 2009.
[28] Murata Y, Yamaguchi S, Kawakami H, Imon Y, Maruyama H, Sakai T, Kazuta T, Ohtake T, Nishimura M, Saida T, Chiba S, Oh-i T, Nakamura S. "Characteristic magnetic resonance imaging findings in Machado-Joseph disease," Arch Neurol., 55(1), pp. 33-37, 1998.
[29] Murata Y, Kawakami H, Yamaguchi S, Nishimura M, Kohriyama T, Ishizaki F, Matsuyama Z, Mimori Y, Nakamura S. "Characteristic magnetic resonance imaging findings in spinocerebellar ataxia 6,". Arch Neurol., 55(10), pp. 1348-1352, 1998.
[30] Morrison MR, Rosenberg RN. "Specific messenger RNA changes in Joseph disease cerebella," Ann Neurol., 14(1), pp. 73-79, 1983.
[31] Wüllner U, Reimold M, Abele M, Bürk K, Minnerop M, Dohmen BM, Machulla HJ, Bares R, Klockgether T. "Dopamine transporter positron emission tomography in spinocerebellar ataxias type 1, 2, 3, and 6," Arch Neurol., 62(8), pp. 1280-1285, 2005.
[32] Soong BW, Liu RS. "Positron emission tomography in asymptomatic gene carriers of Machado-Joseph disease," J Neurol Neurosurg Psychiatry., 64(4), pp. 499-504, 1998.
[33] Marelli C, Cazeneuve C, Brice A, Stevanin G, Dürr A. "Autosomal dominant cerebellar ataxias," Rev Neurol (Paris)., 167(5), pp. 385-400, 2011.
[34] Hara K, Shiga A, Nozaki H, Mitsui J, Takahashi Y, Ishiguro H, Yomono H, Kurisaki H, Goto J, Ikeuchi T, Tsuji S, Nishizawa M, Onodera O. "Total deletion and a missense mutation of ITPR1 in Japanese SCA15 families," Neurology., 71(8), pp. 547-551, 2008.
[35] Duarri A, Jezierska J, Fokkens M, Meijer M, Schelhaas HJ, den Dunnen WF, van Dijk F, Verschuuren-Bemelmans C, Hageman G, van de Vlies P, Küsters B, van de Warrenburg BP, Kremer B, Wijmenga C, Sinke RJ, Swertz MA, Kampinga HH, Boddeke E, Verbeek DS. "Mutations in potassium channel kcnd3 cause spinocerebellar ataxia type 19," Ann Neurol., 72(6), pp. 870-880, 2012.
[36] Delplanque J, Devos D, Huin V, Genet A, Sand O, Moreau C, Goizet C, Charles P, Anheim M, Monin ML, Buée L, Destée A, Grolez G, Delmaire C, Dujardin K, Dellacherie D, Brice A, Stevanin G, Strubi-Vuillaume I, Dürr A, Sablonnière B. "TMEM240 mutations cause spinocerebellar ataxia 21 with mental retardation and severe cognitive impairment," Brain., 137(Pt 10), pp. 2657-2663, 2014.
[37] Lee YC, Durr A, Majczenko K, Huang YH, Liu YC, Lien CC, Tsai PC, Ichikawa Y, Goto J, Monin ML, Li JZ, Chung MY, Mundwiller E, Shakkottai V, Liu TT, Tesson C, Lu YC, Brice A, Tsuji S, Burmeister M, Stevanin G, Soong BW. "Mutations in KCND3 cause spinocerebellar ataxia type 22," Ann Neurol., 72(6), pp. 859-869, 2012.
[38] Bakalkin G, Watanabe H, Jezierska J, Depoorter C, Verschuuren-Bemelmans C, Bazov I, Artemenko KA, Yakovleva T, Dooijes D, Van de Warrenburg BP, Zubarev RA, Kremer B, Knapp PE, Hauser KF, Wijmenga C, Nyberg F, Sinke RJ, Verbeek DS. "Prodynorphin mutations cause the neurodegenerative disorder spinocerebellar ataxia type 23," Am J Hum Genet., 87(5), pp. 593-603, 2010.
[39] Hekman KE, Yu GY, Brown CD, Zhu H, Du X, Gervin K, Undlien DE, Peterson A, Stevanin G, Clark HB, Pulst SM, Bird TD, White KP, Gomez CM. "A conserved eEF2 coding variant in SCA26 leads to loss of translational fidelity and increased susceptibility to proteostatic insult,". Hum Mol Genet., 21(26), pp. 5472-5483, 2012.

Licensed Under Creative Commons Attribution CC BY

[40] Huang L, Chardon JW, Carter MT, Friend KL, Dudding TE, Schwartzentruber J, Zou R, Schofield PW, Douglas S, Bulman DE, Boycott KM. "Missense mutations in ITPR1 cause autosomal dominant congenital nonprogressive spinocerebellar ataxia," Orphanet J Rare Dis., 7, pp. 67, 2012.

[41] Cadieux-Dion M, Turcotte-Gauthier M, Noreau A, Martin C, Meloche C, Gravel M, Drouin CA, Rouleau GA, Nguyen DK, Cossette P. "Expanding the clinical phenotype associated with ELOVL4 mutation: study of a large French-Canadian family with autosomal dominant spinocerebellar ataxia and erythrokeratodermia," JAMA Neurol., 71(4), pp. 470-475, 2014.

[42] Li M, Pang SY, Song Y, Kung MH, Ho SL, Sham PC. "Whole exome sequencing identifies a novel mutation in the transglutaminase 6 gene for spinocerebellar ataxia in a Chinese family," Clin Genet., 83(3), pp.269-273, 2013.

[43] Wang JL, Yang X, Xia K, Hu ZM, Weng L, Jin X, Jiang H, Zhang P, Shen L, Guo JF, Li N, Li YR, Lei LF, Zhou J, Du J, Zhou YF, Pan Q, Wang J, Wang J, Li RQ, Tang BS. "TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing," Brain., 133(Pt 12), pp. 3510-3518, 2010.

[44] Guo YC, Lin JJ, Liao YC, Tsai PC, Lee YC, Soong BW. "Spinocerebellar ataxia 35: novel mutations in TGM6 with clinical and genetic characterization," Neurology., 83(17), pp. 1554-1561, 2014.

[45] Kobayashi H, Abe K, Matsuura T, Ikeda Y, Hitomi T, Akechi Y, Habu T, Liu W, Okuda H, Koizumi A. "Expansion of intronic GGCCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement," Am J Hum Genet., 89(1), pp. 121-130, 2011.

[46] Serrano-Munuera C, Corral-Juan M, Stevanin G, San Nicolás H, Roig C, Corral J, Campos B, de Jorge L, Morcillo-Suárez C, Navarro A, Forlani S, Durr A, Kulisevsky J, Brice A, Sánchez I, Volpini V, Matilla-Dueñas A. "New subtype of spinocerebellar ataxia with altered vertical eye movements mapping to chromosome 1p32," JAMA Neurol., 70(6), pp. 764-771, 2013.

[47] Di Gregorio E, Borroni B, Giorgio E, Lacerenza D, Ferrero M, Lo Buono N, Ragusa N, Mancini C, Gaussen M, Calcia A, Mitro N, Hoxha E, Mura I, Coviello DA, Moon YA, Tesson C, Vaula G, Couarch P, Orsi L, Duregon E, Papotti MG, Deleuze JF, Imbert J, Costanzi C, Padovani A, Giunti P, Maillet-Vioud M, Durr A, Brice A, Tempia F, Funaro A, Boccone L, Caruso D, Stevanin G, Brusco A. "ELOVL5 mutations cause spinocerebellar ataxia 38," Am J Hum Genet., 95(2), pp. 209-217, 2014.

[48] Tsoi H, Yu AC, Chen ZS, Ng NK, Chan AY, Yuen LY, Abrigo JM, Tsang SY, Tsui SK, Tong TM, Lo IF, Lam ST, Mok VC, Wong LK, Ngo JC, Lau KF, Chan TF, Chan HY. "A novel missense mutation in CCDC88C activates the JNK pathway and causes a dominant form of spinocerebellar ataxia," J Med Genet., 51(9), pp. 590-595, 2014.

[49] Coutelier M, Blesneac I, Monteil A, Monin ML, Ando K, Mundwiller E, Brusco A, Le Ber I, Anheim M, Castrioto A, Duyckaerts C, Brice A, Durr A, Lory P, Stevanin G. "A Recurrent Mutation in CACNA1G Alters Cav3.1 T-Type Calcium-Channel Conduction and Causes Autosomal-Dominant Cerebellar Ataxia," Am J Hum Genet., 97(5), pp. 726-737, 2015.

[50] Li SH, McInnis MG, Margolis RL, Antonarakis SE, Ross CA. "Novel triplet repeat containing genes in human brain: cloning, expression, and length polymorphisms," Genomics., 16(3), pp. 572-579, 1993.

[51] Powell A, Chandrasekharan S, Cook-Deegan R. "Spinocerebellar ataxia: patient and health professional perspectives on whether and how patents affect access to clinical genetic testing," Genet Med., 12(4 Suppl), pp. S83-S110, 2010.

[52] Potter NT, Nance MA; "Ataxia Molecular Diagnostic Testing Group. Genetic testing for ataxia in North America," Mol Diagn., 5(2), pp. 91-99, 2000.

[53] Kim JY, Park SS, Joo SI, Kim JM, Jeon BS. "Molecular analysis of Spinocerebellar ataxias in Koreans: frequencies and reference ranges of SCA1, SCA2, SCA3, SCA6, and SCA7," Mol Cells., 12(3), pp. 336-341, 2001.

[54] Gestinari RS, Duarte SF, Pimentel MM, Lima MA. "Genetic polymorphism of MJD1 alleles and molecular analysis of SCA3 patients from Rio de Janeiro, Brazil," Genet Test.; 8(3):281-285, 2004.

[55] Johnson J, Wood N, Giunti P, Houlden H. "Clinical and genetic analysis of spinocerebellar ataxia type 11," Cerebellum., 7(2), pp. 159-164, 2008.

[56] Cao DH, Liu XL, Qiu GB. "Genotyping and prenatal diagnosis of a large spinocerebellar ataxia pedigree in northeastern China," J Genet., 90(2), pp. 369-372, 2011.

[57] Lindblad K, Savontaus ML, Stevanin G, Holmberg M, Digre K, Zander C, Ehrsson H, David G, Benomar A, Nikoskelainen E, Trottier Y, Holmgren G, Ptacek LJ, Anttinen A, Brice A, Schalling M. "An expanded CAG repeat sequence in spinocerebellar ataxia type 7," Genome Res., 6(10), pp. 965-971, 1996.

[58] Martorell L, Pujana MA, Volpini V, Sanchez A, Joven J, Vilella E, Estivill X. "The repeat expansion detection method in the analysis of diseases with CAG/CTG repeat expansion: usefulness and limitations," Hum Mutat., 10(6), pp. 486-488, 1997.

[59] Lindblad K, Lunkes A, Maciel P, Stevanin G, Zander C, Klockgether T, Ratzlaff T, Brice A, Rouleau GA, Hudson T, Auburger G, Schalling M. "Mutation detection in Machado-Joseph disease using repeat expansion detection," Mol Med., 2(1), pp. 77-85, 1996.

[60] Margolis RL. "Dominant spinocerebellar ataxias: a molecular approach to classification, diagnosis, pathogenesis and the future," Expert Rev Mol Diagn., 3(6), pp. 715-732, 2003.

[61] Koutsis G, Pemble S, Sweeney MG, Paudel R, Wood NW, Panas M, Kladi A, Houlden H. "Analysis of spinocerebellar ataxias due to expanded triplet repeats in Greek patients with cerebellar ataxia," J Neurol Sci., 318(1-2), pp. 178-180, 2012.

[62] Dubourg O, Dürr A, Cancel G, Stevanin G, Chneiweiss H, Penet C, Agid Y, Brice A. "Analysis of the SCA1 CAG repeat in a large number of families with dominant ataxia: clinical and molecular correlations," Ann Neurol., 37(2), pp. 176-180, 1995.

[63] Tsai HF, Liu CS, Leu TM, Wen FC, Lin SJ, Liu CC, Yang DK, Li C, Hsieh M. "Analysis of trinucleotide repeats in different SCA loci in spinocerebellar ataxia

Licensed Under Creative Commons Attribution CC BY

patients and in normal population of Taiwan," Acta Neurol Scand., 109(5), pp. 355-360, 2004.
[64] Lau KK, Lam K, Shiu KL, Au KM, Tsoi TH, Chan AY, Li HL, Sheng B. "Clinical features of hereditary spinocerebellar ataxia diagnosed by molecular genetic analysis," Hong Kong Med J., 10(4), pp. 255-259, 2004.
[65] Dorschner MO, Barden D, Stephens K. "Diagnosis of five spinocerebellar ataxia disorders by multiplex amplification and capillary electrophoresis," J Mol Diagn., 4(2), pp. 108-113, 2002.
[66] Lin JX, Ishikawa K, Sakamoto M, Tsunemi T, Ishiguro T, Amino T, Toru S, Kondo I, Mizusawa H. "Direct and accurate measurement of CAG repeat configuration in the ataxin-1 (ATXN-1) gene by "dual fluorescence labeled PCR-restriction fragment length analysis"," J Hum Genet., 53(4), pp. 287-295, 2008.
[67] Biros I I, Forrest SM. "Duplex PCR for Autosomal Dominant Spinocerebellar Ataxia Testing: A Nonradioactive Rapid Screening Method," Mol Diagn., 3(4), pp. 223-227, 1998.
[68] Bauer PO, Kotliarova SE, Matoska V, Musova Z, Hedvicakova P, Boday A, Tomek A, Nukina N, Goetz P. "Fluorescent multiplex PCR--fast method for autosomal dominant spinocerebellar ataxias screening,". Genetika., 41(6), pp. 830-837, 2005.
[69] Hsieh M, Lin SJ, Chen JF, Lin HM, Hsiao KM, Li SY, Li C, Tsai CJ. "Identification of the spinocerebellar ataxia type 7 mutation in Taiwan: application of PCR-based Southern blot," J Neurol., 247(8), pp. 623-629, 2000.
[70] Ishige T, Sawai S, Itoga S, Sato K, Utsuno E, Beppu M, Kanai K, Nishimura M, Matsushita K, Kuwabara S, Nomura F. "Pentanucleotide repeat-primed PCR for genetic diagnosis of spinocerebellar ataxia type 31," J Hum Genet., 57(12), pp. 807-808, 2012.
[71] Krysa W, Rajkiewicz M, Sułek A. "Rapid detection of large expansions in progressive myoclonus epilepsy type 1, myotonic dystrophy type 2 and spinocerebellar ataxia type 8," Neurol Neurochir Pol., 46(2), pp. 113-120, 2012.
[72] Condorelli DF, Trovato-Salinaro A, Spinella F, Valvo S, Saponara R, Giuffrida S. "Rapid touchdown PCR assay for the molecular diagnosis of spinocerebellar ataxia type 2," Int J Clin Lab Res., 28(3), pp. 174-178, 1998.
[73] Shakkottai VG, Fogel BL. "Clinical neurogenetics: autosomal dominant spinocerebellar ataxia," Neurol Clin., 31(4), pp. 987-1007, 2013.
[74] Storey E. "Non-Ataxic Presenting Symptoms of Dominant Ataxias," Cerebellum., 15(1), pp. 26-29, 2016.
[75] Lazarev VF, Benken KA, Semenyuk PI, Sarantseva SV, Bolshakova OI, Mikhaylova ER, Muronetz VI, Guzhova IV, Margulis BA. "GAPDH binders as potential drugs for the therapy of polyglutamine diseases: design of a new screening assay," FEBS Lett., 589(5), pp. 581-587, 2015.
[76] Nakamura K, Mieda T, Suto N, Matsuura S, Hirai H. "Mesenchymal stem cells as a potential therapeutic tool for spinocerebellar ataxia," Cerebellum., 14(2), pp. 165-170, 2015.
[77] Jin JL, Liu Z, Lu ZJ, Guan DN, Wang C, Chen ZB, Zhang J, Zhang WY, Wu JY, Xu Y. "Safety and efficacy of umbilical cord mesenchymal stem cell therapy in hereditary spinocerebellar ataxia," Curr Neurovasc Res., 10(1), pp. 11-20, 2013.
[78] Bavarsad Shahripour R, Harrigan MR, Alexandrov AV. "N-acetylcysteine (NAC) in neurological disorders: mechanisms of action and therapeutic opportunities," Brain Behav., 4(2), pp. 108-122, 2014.
[79] Eldridge R, Iivanainen M, Stern R, Koerber T, Wilder BJ. ""Baltic" myoclonus epilepsy: hereditary disorder of childhood made worse by phenytoin," Lancet., 2(8354), pp. 838-842, 1983.
[80] Tan S, Wang RH, Niu HX, Shi CH, Mao CY, Zhang R, Song B, Sun SL, Liu XJ, Hou HM, Liu YT, Gao Y, Fang H, Kong XD, Xu YM. "Nerve growth factor for the treatment of spinocerebellar ataxia type 3: an open-label study," Chin Med J (Engl)., 128(3), pp. 291-294, 2015.
[81] Bushart DD, Murphy GG, Shakkottai VG. Precision medicine in spinocerebellar ataxias: treatment based on common mechanisms of disease. Ann Transl Med., 4(2), pp. 25, 2016.

Author Profile



Pushpendra Mani Mishra received the Bachelor of Technology in Genetic Engineering from SRM University, Chennai in 2013 and Master of Technology in Biotechnology from Delhi Technological University, Delhi in 2016. He has past research experience from Emerging Pathogen Institute, University of Florida, USA and Institute of Genomics and Integrative Biology-CSIR, New Delhi, India.



Dr. Vimal Kishor Singh, O/I, Stem Cell Research Laboratory, Department of Biotechnology, Delhi Technological University, Delhi, India

Licensed Under Creative Commons Attribution CC BY

Pramana – J. Phys. (2017) 88:22
DOI 10.1007/s12043-016-1330-7

© Indian Academy of Sciences

Minimally coupled scalar field cosmology in anisotropic cosmological model

C P SINGH* and MILAN SRIVASTAVA

Department of Applied Mathematics, Delhi Technological University (Formerly Delhi College of Engineering), Bawana Road, Delhi 110 042, India
*Corresponding author. E-mail: cpsphd@rediffmail.com

MS received 23 January 2016; revised 9 May 2016; accepted 17 June 2016

Abstract. We study a spatially homogeneous and anisotropic cosmological model in the Einstein gravitational theory with a minimally coupled scalar field. We consider a non-interacting combination of scalar field and perfect fluid as the source of matter components which are separately conserved. The dynamics of cosmic scalar fields with a zero rest mass and an exponential potential are studied, respectively. We find that both assumptions of potential along with the average scale factor as an exponential function of scalar field lead to the logarithmic form of scalar field in each case which further gives power-law form of the average scale factor. Using these forms of the average scale factor, exact solutions of the field equations are obtained to the metric functions which represent a power-law and a hybrid expansion, respectively. We find that the zero-rest-mass model expands with decelerated rate and behaves like a stiff matter. In the case of exponential potential function, the model decelerates, accelerates or shows the transition depending on the parameters. The isotropization is observed at late-time evolution of the Universe in the exponential potential model.

Keywords. Cosmology; anisotropic models; scalar field.

PACS Nos 98.80.−Jk; 98.80.−cq; 04.20.−Jb

1. Introduction

Usually, it is assumed that the Universe is filled with a perfect fluid. But observations suggest that cosmological dynamics cannot be fully explained by this standard matter. The observational results lead to the search of some kinds of exotic matter which would generate sufficient negative pressure to drive the late-time cosmic acceleration. One such exotic matter is the scalar field which provides the necessary negative pressure causing acceleration [1,2]. Thus, the cosmological models with scalar field play a vital role in the current modern cosmology to explain the early inflation and the late-time acceleration. The recent discovery of cosmic acceleration [3–7] has stimulated the interest to study cosmological models based on scalar fields. The cosmological models based on scalar fields have been discussed by many researchers for explaining the possible early inflationary scenarios [8] as well as for describing the dark energy [9].

Inflationary models are important in solving several outstanding problems in cosmology like homogeneity, flatness of the observed Universe etc. The dynamics of the evolution of the Universe is often realized by scalar field with a proper scalar potential. The self-interacting potential can act as an effective cosmological constant which drives a period of inflation. It depends on the specific form of the potential as a function of scalar field. Many researchers [10–19] have studied the scalar field cosmology in Friedmann–Robertson–Walker (FRW) model with different forms of scalar potentials like flat, constant and exponential potentials.

A physical entity which provides an accelerated expansion at late-time is commonly known as dark energy. The simplest field having this property is a canonical scalar field with a scalar potential. Nowadays, the scalar field cosmology has acquired a great popularity to analyse the dark energy of the Universe. So far, a large class of scalar field dark energy models have been studied including quintessence [20–23],

a phantom field [24,25], quintom [26,27], k-essence [28], tachyon [29] and so forth.

It is well known that the evolution of the Universe admits a scenario of anisotropic expansion and gains a lot of interest, under the light of the recently announced Planck Probe results [7]. The Bianchi models, which describe homogeneous but anisotropic space–times, have been discussed to explain the significance of anisotropy in the cosmic microwave background (CMB) and large scale structures (LSS) [30–32]. Therefore, motivated by the anomalies found in the CMB anisotropies [33], which violate the statistical isotropy [34], and on the increasing interest on Bianchi cosmologies [35,36], we are interested to investigate the dynamics of a perfect fluid anisotropic Bianchi-I (B-I) cosmological model with a scalar field minimally coupled with gravity. Demianski *et al* [37] have studied the dynamics of anisotropic model filled with scalar field minimally coupled to gravity. Saha and Boyadjiev [38] have considered a self-consistent system of interacting spinor and scalar fields within the framework of a B-I cosmological model filled with perfect fluid. Do *et al* [39] have studied anisotropic power-law inflation for a two-scalar fields model. Sharif and Zubair [40] have studied the behaviour of perfect fluid and massless scalar field for homogeneous and anisotropic B-I Universe model in $f(R, T)$ gravity.

The exact solutions of the field equations play very important roles in cosmology, because they allow us to analyse the qualitative and quantitative behaviour of the Universe as a whole. The number of exact solutions based on scalar field with various forms of scalar potential is limited. It has been observed that in a cosmological model the scalar potential has very important role as it radically affects the cosmological behaviour. Therefore, in this paper, our motivation is to find exact cosmological solutions for a totally homogeneous and anisotropic perfect fluid B-I model with scalar field for various forms of scalar potential. We calculate various observable cosmological parameters like expansion scalar, deceleration parameter (DP), anisotropy parameter, shear scalar, equation of state parameters of scalar field and perfect fluid matter, and the isotropization measure to analyse the dynamics of the evolution of the Universe. We observe that the perfect fluid anisotropic models behave like an isotropic one with the inclusion of exotic matter like scalar field at late-time evolution. The results found in this paper are new and very interesting to discuss the deceleration, acceleration and their transitions. Fadragas *et al* [41] have analysed the anisotropic locally-rotationally-symmetric Bianchi models with scalar field for a wide range of potentials. We extend this work for a totally anisotropic and homogeneous B-I model by assuming two different forms of scalar potential.

The paper is organized as follows: In §2 we present the action of scalar field cosmology in anisotropic space–time metric and its field equations. The various kinematical and physical cosmological parameters are also mentioned. In §3 we discuss the solution of the field equations with two different forms of scalar potential. Section 3 is further divided into two subsections. We obtain the solution of field equations with zero potential and an exponential potential, respectively in §3.1 and 3.2 and discuss the behaviour of each model. Finally, §4 contains the conclusion of the work.

2. Action and the field equations

The action of a Universe constituted with matter and a scalar field minimally coupled to gravity, in the system of units $8\pi G = 1 = c$, reads [20] as

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2}R + \mathcal{L}_\phi + \mathcal{L}_m\right], \tag{1}$$

where g is the determinant of the metric, R is the Ricci scalar, $\mathcal{L}_m$ is the total matter content in the Universe (including cold dark matter (CDM)) and $\mathcal{L}_\phi$ is the scalar field Lagrangian which is defined as

$$\mathcal{L}_\phi = -\frac{1}{2}\nabla^\alpha\phi\nabla_\alpha\phi - V(\phi), \tag{2}$$

where ϕ is the scalar field and $V(\phi)$ is the self-interacting scalar potential. The fluid energy density ρ_m and scalar field ϕ are functions of a time-like coordinate t, and therefore, the variation of S with respect to metric g_{ij} leads to the following gravitational field equations:

$$R_{ij} - \frac{1}{2}g_{ij}R = T^{(m)}_{ij} + T^{(\phi)}_{ij}, \tag{3}$$

where $T^{(m)}_{ij}$ is the energy–momentum tensor of matter which is assumed to be a perfect fluid

$$T^{(m)}_{ij} = (p_m + \rho_m)u_iu_j + p_mg_{ij}, \tag{4}$$

with energy density ρ_m, pressure p_m and its equation of state parameter $w_m = p_m/\rho_m$. In eq. (3), $T^{(\phi)}_{ij}$ is the energy–momentum tensor associated with the scalar field ϕ, which is given by

$$T^{(\phi)}_{ij} = \nabla_i\phi\nabla_j\phi - \left[\frac{1}{2}\nabla^\alpha\phi\nabla_\alpha\phi + V(\phi)\right]g_{ij}. \tag{5}$$

The scalar field itself obeys the wave equation

$$g^{ij}\nabla_i\nabla_j\phi + \frac{\mathrm{d}V(\phi)}{\mathrm{d}\phi} = 0. \tag{6}$$

In what follows we consider a spatially homogeneous and anisotropic B-I space–time described by the line element

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + A^2(t)\mathrm{d}x^2 + B^2(t)\mathrm{d}y^2 + C^2(t)\mathrm{d}z^2, \tag{7}$$

where $A(t)$, $B(t)$ and $C(t)$ are directional scale factors.

With the above assumptions, the Einstein's field equations (3) give the following equations:

$$\frac{\ddot{B}}{B} + \frac{\ddot{C}}{C} + \frac{\dot{B}\dot{C}}{BC} = -(p_m + p_\phi), \tag{8}$$

$$\frac{\ddot{A}}{A} + \frac{\ddot{C}}{C} + \frac{\dot{A}\dot{C}}{AC} = -(p_m + p_\phi), \tag{9}$$

$$\frac{\ddot{A}}{A} + \frac{\ddot{B}}{B} + \frac{\dot{A}\dot{B}}{AB} = -(p_m + p_\phi), \tag{10}$$

$$\frac{\dot{A}\dot{B}}{AB} + \frac{\dot{B}\dot{C}}{BC} + \frac{\dot{C}\dot{A}}{CA} = (\rho_m + \rho_\phi), \tag{11}$$

where an overdot indicates a derivative with respect to time t. Here, ρ_ϕ and p_ϕ are respectively the energy density and pressure for the canonical scalar field, which are given by

$$\rho_\phi = \frac{\dot{\phi}^2}{2} + V(\phi), \quad p_\phi = \frac{\dot{\phi}^2}{2} - V(\phi) \tag{12}$$

and its equation of state parameter is given by

$$\omega_\phi = \frac{p_\phi}{\rho_\phi} = \frac{\dot{\phi}^2 - 2V(\phi)}{\dot{\phi}^2 + 2V(\phi)}. \tag{13}$$

The quintessence cosmological model accommodates a late-time cosmic acceleration when $\omega_\phi < -1/3$ which implies that $\dot{\phi}^2 < V(\phi)$. On the other hand, if the kinetic term of the scalar field is negligible with respect to the potential energy, i.e., $\dot{\phi}^2 \ll 2V(\phi)$, then the EoS is $\omega_\phi = -1$.

It is important to notice that the usual relation $T^{ij}_{\ ;j} = 0$, establishing the conservation laws satisfied by the matter fields, holds true. This follows from the assumption that all matter fields are minimally coupled to the metric g_{ij}, which means that the principle of equivalence is guaranteed. Assuming negligible interaction between the matter and the scalar field, the conservation equation for a perfect fluid leads to

$$\dot{\rho_m} + 3H(1 + \omega_m)\rho_m = 0 \tag{14}$$

and the evolution equation of the scalar field (6) gives

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\mathrm{d}V(\phi)}{\mathrm{d}\phi} = 0 \tag{15}$$

which is equivalently written as $\dot{\rho_\phi} + 3H(1+\omega_\phi)\rho_\phi = 0$.

We now define the average scale factor for the anisotropic model [42,43] as

$$a(t) = (ABC)^{1/3}. \tag{16}$$

The dynamics of the cosmological evolution is characterized by the Hubble parameter. Therefore, in analogy with Hubble parameter in FRW model, we introduce a generalized mean Hubble parameter H as

$$H = \frac{\dot{a}}{a} = \frac{1}{3}\left(\frac{\dot{A}}{A} + \frac{\dot{B}}{B} + \frac{\dot{C}}{C}\right). \tag{17}$$

The directional Hubble parameters which determine the expansion rate of the Universe along x, y and z-axes are given by [42,43]

$$H_1 = \frac{\dot{A}}{A}, \quad H_2 = \frac{\dot{B}}{B}, \quad H_3 = \frac{\dot{C}}{C}. \tag{18}$$

Now, we introduce the kinematical parameters of observational interest in cosmology such as the expansion scalar θ, the deceleration parameter q, the measure of anisotropy parameter A_p and the shear scalar σ^2. We can straightforwardly define these parameters as

$$\theta = u^i_{;\,i} = \left(\frac{\dot{A}}{A} + \frac{\dot{B}}{B} + \frac{\dot{C}}{C}\right), \tag{19}$$

$$q = -1 - \frac{\dot{H}}{H^2}, \tag{20}$$

$$A_p = \frac{1}{3}\sum_{i=1}^{3}\left(\frac{H_i - H}{H}\right)^2, \tag{21}$$

$$\sigma^2 = \frac{1}{2}\sigma_{ij}\sigma^{ij} = \frac{1}{2}\left[\left(\frac{\dot{A}}{A}\right)^2 + \left(\frac{\dot{B}}{B}\right)^2 + \left(\frac{\dot{C}}{C}\right)^2\right] - \frac{\theta^2}{6}. \tag{22}$$

The shear tensor σ^{ij} is defined as

$$\sigma^{ij} = \frac{1}{2}(u_{\mu;\alpha}h^\alpha_\nu + u_{\nu;\alpha}h^\alpha_\mu) - \frac{1}{3}\theta h_{\mu\nu},$$

where $h_{\mu\nu} = g_{\mu\nu} - u_\mu u_\nu$ is the projection tensor. The behaviour of the Universe models is determined by the sign of q. The positive value of deceleration parameter suggests a decelerating model while the negative value indicates inflation.

Using eqs (16)–(22), the field equations (8)–(11) can be rewritten in terms of H, q and σ^2 as

$$\rho_m = 3H^2 - \sigma^2 - \rho_\phi, \tag{23}$$

$$p_m = H^2(2q - 1) - \sigma^2 - p_\phi. \tag{24}$$

In the following section we shall try to find the solutions of the above field equations by assuming suitable physical assumptions.

3. Solution of field equations

Following the method described in [44,45], eqs (8) and (9) give

$$\frac{\dot{A}}{A} - \frac{\dot{B}}{B} = \frac{x_1}{ABC}, \tag{25}$$

where x_1 is a constant of integration. The solution of eq. (25) can be written as

$$\frac{A}{B} = d_1 \exp\left(x_1 \int (ABC)^{-1} \mathrm{d}t\right), \tag{26}$$

where d_1 is another constant of integration. Analogously, from eqs (8) and (10), and eqs (9) and (10), we get

$$\frac{A}{C} = d_2 \exp\left(x_2 \int (ABC)^{-1} \mathrm{d}t\right), \tag{27}$$

$$\frac{B}{C} = d_3 \exp\left(x_3 \int (ABC)^{-1} \mathrm{d}t\right), \tag{28}$$

where d_2, d_3 and x_2, x_3 are constants of integration. Using eq. (16) into eqs (26)–(28), we get the metric functions as

$$A(t) = l_1\, a(t) \exp\left(m_1 \int a^{-3} \mathrm{d}t\right), \tag{29}$$

$$B(t) = l_2\, a(t) \exp\left(m_2 \int a^{-3} \mathrm{d}t\right), \tag{30}$$

$$C(t) = l_3\, a(t) \exp\left(m_3 \int a^{-3} \mathrm{d}t\right), \tag{31}$$

where

$$l_1 = \sqrt[3]{d_1 d_2}, \quad l_2 = \sqrt[3]{d_1^{-1} d_3}, \quad l_3 = \sqrt[3]{d_2 d_3^{-1}}$$

and

$$m_1 = \frac{x_1 + x_2}{3}, \quad m_2 = \frac{x_3 - x_1}{3}, \quad m_3 = -\frac{x_3 + x_2}{3}.$$

The constants l_1, l_2, l_3, and m_1, m_2, m_3 satisfy the following relations:

$$l_1 l_2 l_3 = 1 \quad \text{and} \quad m_1 + m_2 + m_3 = 0. \tag{32}$$

The unknown quantities of the problem are A, B, C, p_m, ρ_m, ϕ and $V(\phi)$ whereas we have only five equations available, namely eqs (8)–(11) and (15). Therefore, to solve the system of differential equations we need to assume two more relations among the unknown quantities. Thus, for any arbitrary average scale factor a and scalar field potential $V(\phi)$, eq. (15) gives $\phi(t)$ and eqs (29)–(31) give $A(t)$, $B(t)$ and $C(t)$. In this paper we assume a relationship between the average scale factor a and scalar field ϕ of the form [46,47]

$$a = a_0 \mathrm{e}^{\alpha\phi(t)}, \tag{33}$$

where a_0 represents the average scale factor at the present time and α is an arbitrary positive constant. Thus, the Hubble parameter simply gives

$$H = \alpha\dot{\phi}(t). \tag{34}$$

To illustrate our analysis, we assume a specific form of $V(\phi)$ in the following subsection. In the literature, due to the unknown nature of dark energy, there are various forms of this potential (for a detailed review, see [48–55]) which describe the physical features of the scalar field cosmology. Therefore, as long as potential $V(\phi)$ is given, we can solve other geometrical and physical parameters correspondingly.

3.1 *Solution with zero potential*

In this section, we consider the case where the scalar potential is zero [53,54], that is,

$$V(\phi) = 0. \tag{35}$$

Using eqs (34) and (35), eq. (15) reduces to

$$\ddot{\phi} + 3\alpha\dot{\phi}^2 = 0. \tag{36}$$

The solution of eq. (36) is given by

$$\phi = \frac{1}{3\alpha} \ln[3\alpha(c_1 t + t_0)], \tag{37}$$

where c_1 and t_0 are positive constants of integration. We observe that at the beginning, i.e., at $t = -(t_0/c_1)$, $\phi \to -\infty$. It increases with time and tends to $+\infty$ at $t \to \infty$. Thus, the kinetic energy vanishes at the end of the evolution (an infinite expansion).

Using eq. (37) into eq. (33), the average scale factor in terms of t for an expanding Universe is given by

$$a = a_0[3\alpha(c_1 t + t_0)]^{1/3}. \tag{38}$$

This is of the power-law form $a \propto t^r$ which represents a generalized inflation for $r > 1$. But, in our case we have $r = 1/3 < 1$. Therefore, the model will expand with decelerated rate.

Using eq. (38) into the metric functions (29)–(31), we get

$$A(t) = a_0 l_1 [3\alpha(c_1 t + t_0)]^{\frac{1}{3}\left(1 + \frac{m_1}{\alpha c_1 a_0^3}\right)}, \tag{39}$$

$$B(t) = a_0 l_2 [3\alpha(c_1 t + t_0)]^{\frac{1}{3}\left(1 + \frac{m_2}{\alpha c_1 a_0^3}\right)}, \tag{40}$$

$$C(t) = a_0 l_3 [3\alpha(c_1 t + t_0)]^{\frac{1}{3}\left(1+\frac{m_3}{\alpha c_1 a_0^3}\right)}. \quad (41)$$

Equations (39)–(41) show that the directional scale factors have power-law expansion form. These three spatial scale factors are zero at $t = -t_0/c_1$ and all tend to infinity at $t \to \infty$. The model has a point singularity at $t = -t_0/c_1$. The directional Hubble parameters have the expressions:

$$H_i = \alpha c_1 \left(1 + \frac{m_i}{\alpha c_1 a_0^3}\right)[3\alpha(c_1 t + t_0)]^{-1}, \quad i = 1, 2, 3. \quad (42)$$

Using the constraint $m_1 + m_2 + m_3 = 0$, the generalized mean Hubble parameter is given by

$$H = \frac{c_1}{3(c_1 t + t_0)}. \quad (43)$$

As H is a function of time, the model is not a steady-state model. The DP, defined in eq. (20), gives

$$q = 2, \quad (44)$$

which shows that the model expands with decelerated rate. From eqs (19), (21) and (22), we have the following values of physical quantities:

$$\theta = c_1 (c_1 t + t_0)^{-1}, \quad (45)$$

$$A_p = \frac{2M}{3\alpha^2 c_1^2 a_0^6}, \quad (46)$$

$$\sigma^2 = \frac{M}{9\alpha^2 a_0^6} \frac{1}{(c_1 t + t_0)^2}, \quad (47)$$

where

$$M = \frac{m_1^2 + m_2^2 + m_3^2}{2}.$$

The values of ρ_ϕ and p_ϕ are given by

$$\rho_\phi = \frac{c_1^2}{2}[3\alpha(c_1 t + t_0)]^{-2} = p_\phi, \quad (48)$$

which gives $w_\phi = 1$ throughout the evolution and this is the case of stiff-matter state. We observe that θ, σ, ρ_ϕ and p_ϕ are decreasing functions of time which have infinite value at $t = -t_0/c_1$ but tend to zero in late-time evolution. The anisotropy parameter is constant, which shows that the nature of the model is always anisotropic throughout the evolution. Also, $\lim_{t\to\infty} \frac{\sigma}{\theta} = \text{const.}$, i.e., the measure of shear scalar to expansion rate is constant and continues throughout the evolution, which shows that the shear does not tend to zero faster than the expansion scalar and hence the model has anisotropy behaviour.

Now, eqs (23) and (24) give

$$\rho_m = \left(3\alpha^2 c_1^2 - \frac{M}{a_0^6} - \frac{c_1^2}{2}\right)[3\alpha(c_1 t + t_0)]^{-2} = p_m. \quad (49)$$

For energy density to be positive we must have the positive value within the first bracket, i.e., $6\alpha^2 c_1^2 a_0^6 - c_1^2 a_0^6 - 2M > 0$. From eq. (49), we get $w_m = 1$, which is again the case of stiff-matter. The effective density ($\rho_{\text{eff}} = \rho_m + \rho_\phi$) and pressure ($p_{\text{eff}} = p_m + p_\phi$) are respectively given by

$$\rho_{\text{eff}} = \left(3\alpha^2 c_1^2 - \frac{M}{a_0^6}\right)[3\alpha(c_1 t + t_0)]^{-2} = p_{\text{eff}}. \quad (50)$$

From eq. (50) we find that $w_{\text{eff}} = 1$. Thus, the Universe is filled with stiff-matter in the presence of scalar field with zero potential and gives decelerating Universe as $q > 0$.

3.2 *Solution with exponential potential*

Let us assume the exponential potential of the form [10,48,50]

$$V(\phi) = V_0 e^{-\lambda\phi}, \quad (51)$$

where V_0 and λ are non-negative constants.

Using (34) and (51), eq. (15) reduces to

$$\ddot{\phi} + 3\alpha\dot{\phi}^2 - V_0\lambda e^{-\lambda\phi} = 0. \quad (52)$$

The solution of eq. (52) is given by

$$\phi = \frac{2}{\lambda}\left[\ln\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)\right], \quad (53)$$

where $c_2 > 0$ is a constant of integration and $D = 2V_0\lambda/(6\alpha - \lambda)$. The real solution of ϕ exists provided $D > 0$, i.e., $0 < \lambda < 6\alpha$. We shall choose the positive sign within the bracket, without loosing any generality to obtain an expanding model. We find that ϕ is time-dependent and is increasing function of cosmic time. Therefore, during the evolution the scalar field is growing and hence kinetic energy vanishes at late-time evolution.

Substituting eq. (53) into eq. (33), the solution of the average scale factor for an expanding Universe is given by

$$a = a_0\left[\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right]^{2\alpha/\lambda}. \quad (54)$$

In this case, we again get the power-law form $a \propto t^r$. The model will accelerate or decelerate according to

$0 < \lambda \leq 2\alpha$ or $2\alpha < \lambda < 6\alpha$. From eqs (51) and (53), the potential turns out to be

$$V(\phi) = V_0 \left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-2}, \tag{55}$$

which shows that V is a decreasing function of time and vanishes as $t \to \infty$. By use of eq. (54) in eqs (29)–(31), the spatial scale factors can be obtained as

$$A(t) = a_0 l_1 \left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{2\alpha/\lambda} \times \exp\left[-\frac{2m_1}{a_0^3\sqrt{D}(6\alpha-\lambda)} \frac{1}{\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{\left(\frac{6\alpha-\lambda}{\lambda}\right)}}\right], \tag{56}$$

$$B(t) = a_0 l_2 \left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{2\alpha/\lambda} \times \exp\left[-\frac{2m_2}{a_0^3\sqrt{D}(6\alpha-\lambda)} \frac{1}{\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{\left(\frac{6\alpha-\lambda}{\lambda}\right)}}\right], \tag{57}$$

$$C(t) = a_0 l_3 \left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{2\alpha/\lambda} \times \exp\left[-\frac{2m_3}{a_0^3\sqrt{D}(6\alpha-\lambda)} \frac{1}{\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{\left(\frac{6\alpha-\lambda}{\lambda}\right)}}\right]. \tag{58}$$

Equations (56)–(58) show that the scale factors have hybrid (a combination of power-law and exponential) type expansion. As we know the power-law behaviour dominates in the early phase of the cosmic evolution whereas the exponential factor dominates at late phase. Therefore, this form of scale factor describes both types of expansion depending on the dominating factor. The model has a point singularity at $t = -c_2/\sqrt{D}$. This solution describes an evolution from a point singularity to an infinite expansion. The directional Hubble parameters in this model are given by

$$H_i = \alpha\sqrt{D}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-1} + \frac{m_i}{a_0^3}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-6\alpha/\lambda}, \quad i = 1, 2, 3. \tag{59}$$

Now, the generalized mean Hubble parameter is obtained as

$$H = \alpha\sqrt{D}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-1}, \quad m_1+m_2+m_3 = 0. \tag{60}$$

The deceleration parameter (DP) is given by

$$q = \frac{\lambda}{2\alpha} - 1. \tag{61}$$

Equation (61) clearly shows that q is a constant and therefore, its nature ($q < 0$, $q = 0$, or $q > 0$) depends on the values of α and λ. Therefore, for an accelerating Universe where $-1 < q < 0$, we must have $0 < \lambda < 2\alpha$, for marginal inflation where $q = 0$, we have $\lambda = 2\alpha$ and for a decelerating Universe where $q > 0$, we must have the constraints $2\alpha < \lambda < 6\alpha$. It is to be noted that $\lambda > 6\alpha$ gives the imaginary value.

The scalar field density and pressure are respectively given by

$$\rho_\phi = \frac{6\alpha V_0}{(6\alpha - \lambda)}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-2}, \tag{62}$$

$$p_\phi = \frac{2V_0(\lambda - 3\alpha)}{(6\alpha - \lambda)}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-2}. \tag{63}$$

We observe that ρ_ϕ is always positive as $\lambda < 6\alpha$, and both ρ_ϕ and p_ϕ are decreasing functions of time. The corresponding EoS parameter is given by

$$w_\phi = \frac{\lambda}{3\alpha} - 1 = \frac{2q-1}{3}. \tag{64}$$

As we know, the EoS parameter to a quintessence region is $-1 < w_\phi < -(1/3)$, and this can be achieved when $0 < \lambda < 2\alpha$. Also, $2\alpha < \lambda < 6\alpha$ gives the EoS parameter $-(1/3) < w_\phi < 1$.

The parameters θ, A_p and σ^2 are respectively obtained as

$$\theta = 3\alpha\sqrt{D}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-1}, \tag{65}$$

$$A_p = \frac{2M}{3\alpha^2 D a_0^6}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{(-12\alpha/\lambda)+2}, \tag{66}$$

$$\sigma^2 = \frac{M}{a_0^6}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}. \tag{67}$$

We observe that θ, A_p and σ^2 are infinite at the point of singularity but they decrease with time and all tend to zero in late-time evolution. Thus, the Universe is expanding with the increase of time but the rate of expansion, measure of anisotropy and shear scalar decrease to zero and become isotropic in late time. It is also observed that $\lim_{t\to\infty}(\sigma/\theta) = 0$ for $\lambda < 6\alpha$ which shows that the model approaches isotropy in late time. Therefore, the anisotropy of the Universe damp out during the course of evolution which is consistent with the present observation. Now, eqs (23) and (24) give

$$\rho_m = \frac{6\alpha V_0(\alpha\lambda - 1)}{(6\alpha - \lambda)}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-2} - \frac{M}{a_0^6}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}, \tag{68}$$

$$p_m = \frac{2V_0(\alpha\lambda - 1)(\lambda - 3\alpha)}{(6\alpha - \lambda)}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-2} - \frac{M}{a_0^6}\left(\frac{\lambda}{2}c_2 + \frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}. \tag{69}$$

The condition $\rho_m \geq 0$ gives

$$t \geq \left(\frac{2c_3}{\lambda\sqrt{D}} - \frac{c_2}{\sqrt{D}}\right),$$

where

$$c_3 = \left[\frac{M(6\alpha - \lambda)}{6\alpha V_0 a_0^6(\alpha\lambda - 1)}\right]^{\lambda/2(6\alpha-\lambda)}$$

with $\alpha\lambda \geq 1$.

The corresponding perfect fluid EoS parameter w_m is given by

$$w_m = \frac{\frac{2V_0(\alpha\lambda-1)(\lambda-3\alpha)\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-2}}{(6\alpha-\lambda)} - \frac{M\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}}{a_0^6}}{\frac{6\alpha V_0(\alpha\lambda-1)\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-2}}{(6\alpha-\lambda)} - \frac{M\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}}{a_0^6}}. \tag{70}$$

Now with the EoS parameter (70) it is obvious that, as long as both ρ_m and p_m are positive, $0 < w_m < 1$ only when $3\alpha < \lambda < 6\alpha$ and as p_m is negative, $-1 \leq w_m \leq 0$ only when $0 < \lambda \leq 3\alpha$.

We can also calculate $\rho_{\text{eff}} = \rho_\phi + \rho_m$ and $p_{\text{eff}} = p_\phi + p_m$ but we avoid here to write the expressions of these quantities. The effective EoS parameter is calculated as

$$w_{\text{eff}} = \frac{\frac{2V_0\alpha\lambda(\lambda-3\alpha)\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-2}}{(6\alpha-\lambda)} - \frac{M\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}}{a_0^6}}{\frac{6\alpha^2\lambda V_0\left(\frac{\lambda}{2}c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-2}}{(6\alpha-\lambda)} - \frac{M\left(\lambda 2c_2+\frac{\lambda\sqrt{D}}{2}t\right)^{-12\alpha/\lambda}}{a_0^6}}. \tag{71}$$

In this case we can give similar interpretation. As long as both ρ_{eff} and p_{eff} are positive, $0 < w_{\text{eff}} < 1$ only when $3\alpha < \lambda < 6\alpha$ and as p_{eff} is negative, $-1 \leq w_{\text{eff}} \leq 0$ only when $0 < \lambda \leq 3\alpha$.

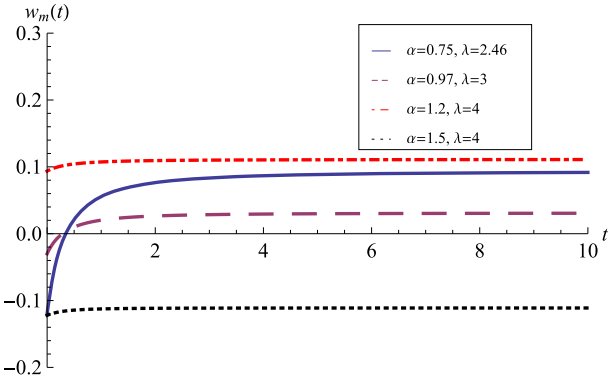


Figure 1. w_m vs. t for $c_2 = 1$, $V_0 = 1$ and some values of λ and α.

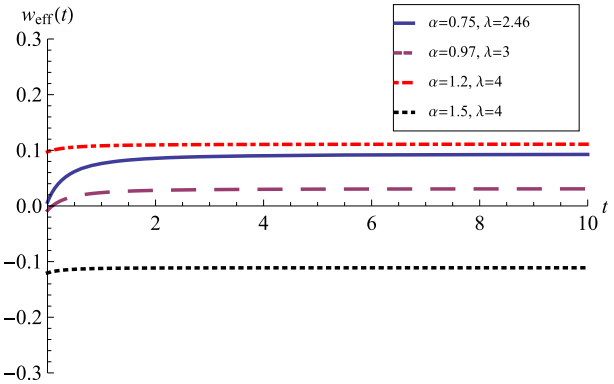


Figure 2. w_{eff} vs. t for $c_2 = 1$, $V_0 = 1$ and some values of λ and α.

Figures 1 and 2 plot the graphs of w_m and w_{eff}, respectively with respect to time t for some values of α and λ satisfying the above-mentioned constraints. It is to be noted that one can take those values of α and λ which make ρ_m and ρ_{eff} positive. We observe that for some values of parameters, e.g., $\alpha = 0.97$ and $\lambda = 3$, both w_m and w_{eff} show a transition from a negative value to a positive one. Therefore, the model shows transition from early inflationary phase to decelerated phase. There are some values like $\alpha = 0.75$ and $\lambda = 2.46$, where w_m shows transition from negative to positive but w_{eff} is always positive. However, there are some values of α and λ for which w_m and w_{eff} are only positive or negative. Thus, depending on the choices of arbitrary values of α and λ under the constraints we can observe any one of the behaviour in EoS parameters: w_m and w_{eff} may be only positive or negative throughout the evolution or may vary from negative to positive.

4. Conclusion

We have presented a detailed discussion on the evolution of a homogeneous and anisotropic B-I Universe with a perfect fluid and a scalar field. We have studied two types of models, namely, zero potential and an exponential potential. Exact solutions of the field equations for both the models have been obtained by assuming the average scale factor as an exponential function of the scalar field. Explicit forms of average scale factor (also, the spatial scale factors), energy density and pressure of the scalar field and the perfect fluid, and their respective EoS parameters have been obtained in each model. Other geometrical and physical parameters like expansion scalar, deceleration parameter, anisotropy and shear scalar have been calculated and their physical interpretations have been discussed. We have observed that these geometries are very interesting for the observed Universe isotropy as we find the late-time isotropized solutions. The solutions obtained in both models satisfy the energy conservation equation (14). We, herewith, summarize the result of both these models as follows.

In the first model, we have considered a massless scalar field by assuming the average scale factor as an exponential function of the scalar field. We have solved the field equations and found solutions for various cosmological parameters. In this case, the average scale factor as well as the directional scale factors have power-law expansion which expand with decelerated rate. The model has a point singularity at $t = -t_0/c_1$. At the beginning of the evolution, i.e., at $t = -t_0/c_1$, the scalar field $\phi \to -\infty$. During the evolution, ϕ increases and at the end of the evolution (an infinite expansion) when $t \to \infty$, it tends to $+\infty$. The kinetic energy vanishes at $t \to \infty$. The positive constant value of DP shows that the model expands but with decelerated rate. The nature of the model is always anisotropic throughout the evolution because of the constant behaviour of anisotropy parameter. All the physical parameters like energy density, pressure, expansion scalar and shear scalar are decreasing functions of time. The measure of shear scalar to expansion scale is constant throughout the evolution. Thus, the shear scalar does not tend faster than the expansion scalar. The EoS parameters of the scalar field and the perfect fluid come out to be one. Hence, we have found that the anisotropic perfect fluid model behaves like stiff-matter in the presence of scalar field.

In the second model, we have considered an exponential potential, i.e., $V(\phi) = V_0 e^{-\lambda\phi}$. We have obtained the logarithmic form of the scalar field and power-law form of the average scale factor. The directional scale factors have hybrid-type expansion which is a combination of power-law and exponential forms.

This form of scale factor describes both types of expansion depending on the dominating factor. The model has a point singularity at $t = -c_2/\sqrt{D}$. It has been observed that at the beginning of the evolution, i.e., at $t = -c_2/\sqrt{D}$, ϕ tends to $-\infty$ while $V(\phi)$ tends to $+\infty$. During the evolution ϕ increases whereas $V(\phi)$ decreases and at late-time evolution ϕ tends to $+\infty$ while $V(\phi)$ vanishes. The kinetic energy and potential energy tend to $+\infty$ at the beginning but both vanish at the end of evolution. This means that the Universe is born from the singularity at $t = -c_2/\sqrt{D}$. The DP is constant whose nature depends on the values of α and λ. The model shows the quintessence region for $0 < \lambda < 2\alpha$. The model has marginal inflation at $\lambda = 2\alpha$. The model decelerates for $2\alpha < \lambda < 6\alpha$. Physical quantities like θ, A_p, σ^2, ρ_m, p_m, ρ_ϕ, p_ϕ are decreasing functions of time and tend to zero in late-time evolution. The model approaches isotropy in late-time. The EoS parameter of matter and the effective EoS parameter of matter plus scalar field are time dependent which show that the deceleration or acceleration or transition from one phase to other depends on the constraint $0 < \lambda \leq 3\alpha$ or $3\alpha < \lambda < 6\alpha$ (shown in figures 1 and 2). We find that the physically valid range for λ is $0 < \lambda < 6\alpha$, where α is any positive constant.

Acknowledgements

The authors are thankful to the referee for the constructive suggestions to improve the manuscript. MS also expresses her thanks to UGC, India for providing financial support to carry out the research under Junior Research Fellowship (JRF).

References

[1] R R Caldwell, R Dave and P J Steinhardt, *Phys. Rev. Lett.* **80**, 1582 (1998)
[2] V Sahni, *Lect. Notes Phys.* **653**, 141 (2004)
[3] A G Riess *et al*, *Astron. J.* **116**, 1009 (1998)
[4] S Perlmutter *et al*, *Astrophys. J.* **517**, 565 (1999)
[5] N Suzuki *et al*, *Astrophys. J.* **746**, 85 (2012)
[6] G Hinshaw *et al*, *Astrophys. J. Suppl.* **208**, 19 (2013)
[7] Planck Collaboration: P A R Ade *et al*, *Astron. Astrophys.* **571**, A16 (2014)
[8] A H Guth, *Phys. Rev. D* **23**, 347 (1981)
[9] J A Frieman and I Waga, *Phys. Rev. D* **57**, 4642 (1998)
[10] A B Burd and J D Barrow, *Nucl. Phys. B* **308**, 929 (1988)
[11] J D Barrow, *Phys. Lett. B* **187**, 12 (1987); *Phys. Rev. D* **49**, 3055 (1994)
[12] G F R Ellis and M S Madsen, *Class. Quantum Grav.* **8**, 667 (1991)
[13] J D Barrow and P Saich, *Class. Quantum Grav.* **10**, 279 (1993)
[14] J D Barrow and P Parsons, *Phys. Rev. D* **52**, 5576 (1995)
[15] P Parsons and J D Barrow, *Phys. Rev. D* **51**, 6757 (1995)
[16] A A Coley, J Ibanez and R J Van den Hoogen, *J. Math. Phys.* **38**, 5256 (1997)
[17] E J Copeland, A R Liddle and D Wands, *Phys. Rev. D* **57**, 4686 (1998)
[18] I P C Heard and D Wands, *Class. Quantum Grav.* **19**, 5435 (2002)
[19] E J Copeland, S Mizuno and M Shaeri, *Phys. Rev. D* **79**, 103515 (2009)
[20] B Ratra and P J E Peebles, *Phys. Rev. D* **37**, 3406 (1988)
[21] A R Liddle and R J Scherrer, *Phys. Rev. D* **59**, 023509 (1999)
[22] E J Copeland, M Sami and S Tsujikawa, *Int. J. Mod. Phys. D* **15**, 1753 (2006)
[23] S Dutta, E N Saridakis and R J Scherrer, *Phys. Rev. D* **79**, 103005 (2009)
[24] R R Caldwell, *Phys. Lett. B* **545**, 23 (2002)
[25] S Nojiri and S D Odintsov, *Phys. Lett. B* **562**, 147 (2003)
[26] Z-K Guo, Y-S Piao, X Zhang and Y-Z Zhang, *Phys. Lett. B* **608**, 177 (2005)
[27] Y-F Cai, E N Saridakis, M R Setare and J-Q Xia, *Phys. Rep.* **493**, 1 (2010)
[28] C Armendariz-Picon, V Mukhanov and P J Steinhardt, *Phys. Rev. D* **63**, 103510 (2001)
[29] A Sen, *J. High Energy Phys.* **0207**, 065 (2002)
[30] SDSS Collaboration: M Tegmark *et al*, *Phys. Rev. D* **69**, 103501 (2004)
[31] WMAP Collaboration: J Donkley *et al*, *Astrophys. J. Suppl.* **180**, 306 (2009)
[32] WMAP Collaboration: E Komatsu *et al*, *Astrophys. J. Suppl.* **192**, 18 (2011)
[33] C J Copi, D Huterer and G D Starkman, *Phys. Rev. D* **70**, 043515 (2004)
[34] A Bernui, B Mota, M J Reboucas and R Tavakol, *Int. J. Mod. Phys. D* **16**, 411 (2007)
[35] G F R Ellis and M A H MacCallum, *Commun. Math. Phys.* **12**, 108 (1969)
[36] G F R Ellis, *Gen. Rel. Gravit.* **38**, 1003 (2006)
[37] M Demianski, R De Ritis, C Rubano and P Scudellaro, *Phys. Rev. D* **46**, 1391 (1992)
[38] B Saha and T Boyadjiev, *Phys. Rev. D* **69**, 124010 (2004)
[39] T Q Do, W F Kao and I-C Lin, *Phys. Rev. D* **83**, 123002 (2011)
[40] M Sharif and M Zubair, *J. Phys. Soc. Jpn* **81**, 114005 (2012)
[41] C R Fadragas, G Leon and E N Saridakis, *Class. Quantum Grav.* **31**, 075018 (2014), arXiv:gr-qc/1308.1658
[42] Ø Grøn, *Phys. Rev. D* **32**, 2522 (1985)
[43] Ø Grøn, *Astrophys. Space Sci.* **173**, 191 (1990)
[44] B Saha and V Rikhvitsky, *Physica D* **219**, 168 (2006)

[45] S Kumar and C P Singh, *Astrophys. Space Sci.* **312**, 57 (2007)
[46] S Sen and N Banerjee, *Astrophys. Space Sci.* **254**, 133 (1997)
[47] C P Singh and V Singh, *Mod. Phys. Lett. A* **26**, 1495 (2011)
[48] L P Chimento, *Class. Quantum Grav.* **15**, 965 (1998)
[49] J W Van Holten, *Mod. Phys. Lett. A* **17**, 1383 (2002)
[50] J G Russo, *Phys. Lett. B* **600**, 185 (2004), arXiv: hep-th/0403010
[51] E J Copeland, M Sami and S Tsujikawa, *Int. J. Mod. Phys. D* **15**, 1753 (2006)
[52] R R Caldwell and M Kamionkowski, *Annu. Rev. Nucl. Part. Sci.* **59**, 397 (2009)
[53] S V Sushkov, *Phys. Rev. D* **80**, 103505 (2009)
[54] E N Saridakis and S V Sushkov, *Phys. Rev. D* **81**, 083510 (2010), arXiv:gr-qc/1002.3478
[55] V Singh and C P Singh, *Astrophys. Space Sci.* **355**, 2183 (2014)

Volume 4, Issue 4(2016) 692-697 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

Numerical Simulation of NACA 2415 Airfoil at Different Low Reynolds Numbers and Angles of Attack

Aakash Srivastava, Vatsalya Tiwari, Raj Kumar Singh
Department of Mechanical Engineering, Delhi Technological University Bawana Road, Delhi

Article Info
Article history:
Received 25 September 2016
Received in revised form
20 October 2016
Accepted 28 November 2016
Available online 15 December 2016

Keywords: Computational Fluid Dynamics (CFD), NACA2415, FLUENT, Spalart-Allmaras, Validation of results Effect of Reynolds number and Angle of Attack.

Abstract
This project simulates NACA2415 airfoil on ANSYS Workbench and ANSYS FLUENT at low Reynolds numbers at different angles of attack. This is a 2-D simulation andSpalart-Allmaras is the preferred turbulentmodel solver for this process, it yielded more results closer to experimental results when compared against K-epsilon and other turbulent models.Contours of Pressure and Velocity are presented in this paper with their inferences discussed while Plots of Coefficient of Pressure (C_P) about the chord lengths along the airfoil and Coefficient of Lift (C_L) are plotted to compare the CFD and the experimental results. Effect of Reynolds number and Angle of Attack is thus studied and investigated.

1. Introduction

Ever since the emergence of powerful computers and advanced numerical techniques and algorithms, computational fluid dynamics (CFD) has revolutionized aerospace industries all over the world. While experimental testing will always remain an integral part of the design, CFD is reducing the dependencies on experimental testing and has emerged as one of the important design tools for aerospace vehicle design [1]. Being cost-efficient, reliable and productive, CFD finds its usage in solving complicated problems, where testing experimentally in say, wind tunnels is difficult and expensive.

An **airfoil** or **aerofoil** is the shape of a wing or blade (of a propeller, rotor, or turbine) or sail as seen in cross-section. An airfoil-shaped body moved through a fluid produces an aerodynamic force. The component of this force perpendicular to the direction of motion is called lift. The component parallel to the direction of motion is called drag. Subsonic flight airfoils have a characteristic shape with a rounded leading edge, followed by a sharp trailing edge, often with a symmetric curvature of upper and lower surfaces. The lift on an airfoil is primarily the result of its angle of attack and shape. When oriented at a suitable angle, the airfoil deflects the oncoming air, resulting in a force on the airfoil in the direction opposite to the deflection. This force is known as aerodynamic force and can be resolved into two components: lift and drag [2].

2. Nomenclature

The shape of an airfoil is depicted in the Figure.1. The **leading edge** is the point at the front of the airfoil that has maximum curvature. The **trailing edge** is defined similarly as the point of maximum curvature at the rear of the airfoil. The **chord line** is a straight line connecting the leading and trailing edges of the airfoil. The *chord length* or simply **chord** is the length of the chord line and is the characteristic dimension of the airfoil section [3].

Hasanuzzaman Md, Mashud, Md. (2013), used FLUENT to see effect of Reynolds number on aerodynamic characteristics of NACA2415 airfoil with flap and they concluded flow separation can be controlled by using flap at maximum camber position [4].

Ghods, M. (2001) conducting an experiment to introduced the basic theory of wing and provide an introduction on wind tunnel testing involving NACA2415. Lift increases as the angle of attack increases, maximum liftis generated [5].

Ismail, Md. (2013), did investigation on the Aerodynamic Characteristics of NACA2415 airfoil at Reynolds number of 200000 and calculated Cd, Cl, Cp and discussed the formation of Laminar Separation Bubble [6].

Luis Velazquez-Araque and Jiří Nožička (2014), by means of CFD using k-epsilon model obtained lift, drag and pitching moment coefficients and also the flow field of air past airfoil 2415-3S.[7]

The aim of this research was to perform parametric investigate the aerodynamics characteristics of a NACA2415 airfoil system.

Computations were performed for various configurations of the system where each configuration represents a different Re number and angle of attack. The type of flow observed was subsonic and turbulent. Numerical simulations were focused on visualizing and analyzing pressure, velocity and lift of the airfoil system in steady state. Numerical simulations were performed using the hybrid grid Reynolds Averaged Navier-Stokes solver, ANSYS FLUENTTM. It is necessary to generate the computational grids and choose a computational turbulence model that could capture the complex flow phenomena around the airfoil. Initial studies were primarily focused on choosing the best turbulence model for the numerical simulations of the airfoil system. The turbulence model is critical to the computational solution and care must be taken to choose a model that can accurately compute and resolve the complex flow around the airfoil. Ideal and most accurate turbulence model was found out to be Spalart-Allmaras turbulent model. Other models, namely, K-Epsilon and K-Omega were also tested against it. Turbulence model validation was performed by computationally analyzing the NACA 2415 airfoil section and comparing the lift, drag, and pressure coefficient results with experimental data.

***Corresponding Author,**
E-mail address: rksingh.dtu@gmail.com; cheeku1612@gmail.com, atsalya.tiwari.vns@gmail.com,
All rights reserved: http://www.ijari.org

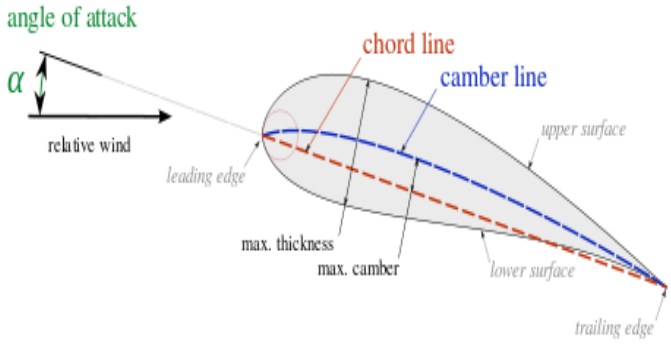


Fig.1 Nomenclature and shape of an Airfoil.[2]

2.1 Model

ANSYS FLUENT software could provide many turbulence models including Spalart-Allmaras model, k-model and the Reynolds stress model and K-Omega model, etc. Spalart-Allmaras turbulence model was put forward by S. Allmaras in 1992, which is a simple-equation turbulence model through solving the turbulence viscosity of the transport equations. It is the preferred model for the numerical simulations in this project. This turbulence model is a single-equation turbulence model applied to the turbulent flow along the solid wall boundary including the laminar flow, which has a smaller calculation, a good stability and a high precision.

The one-equation model is given by the following equation: [8]

$$\frac{\partial \hat{\nu}}{\partial t}+u_j\frac{\partial \hat{\nu}}{\partial x_j}=c_{b1}(1-f_{t2})\hat{S}\hat{\nu}-\left[c_{w1}f_w-\frac{c_{b1}}{\kappa^2}f_{t2}\right]\left(\frac{\hat{\nu}}{d}\right)^2+\frac{1}{\sigma}\left[\frac{\partial}{\partial x_j}\left((\nu+\hat{\nu})\frac{\partial \hat{\nu}}{\partial x_j}\right)+c_{b2}\frac{\partial \hat{\nu}}{\partial x_i}\frac{\partial \hat{\nu}}{\partial x_i}\right]$$

2.2. Grid Generation

The numerical simulation of airfoil NACA 2415 is performed using ICEM software to build the calculated model whose chord length is taken as unity. Computational solutions are sensitive to the discretization of the computational domain. For the present 2D computations, grids, as shown in Fig. 3, were generated with structured cells consisting of very fine quadrilateral elements in the boundary layer region of the airfoil and less dense cells consisting of quadrilateral elements for the rest of the domain. The computational domain is the geometrical region which bounds the numerical simulation. The distance of the far-field region from the surface of the airfoil must be such that the effects of the flow at the far-field region do not have an impact on the near-field solution. The computational domain area is determined by the far-field distance given in terms of the chord length of the airfoil under consideration. Figures 2.a and 2.b are views of grid generated for this problem, of airfoil and of entire domain respectively. Computational domain area for the present grids was defined by placing the far-field boundaries at 10 times the system chord length. The total grid consists of 40200 cells, 40602 nodes and 80802 faces.

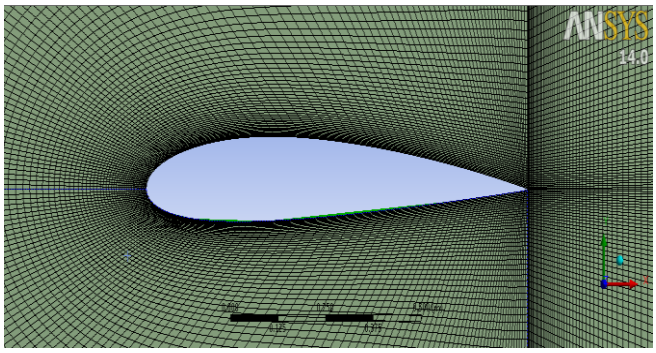


Fig. 2(a) Zoomed-in views of the airfoil mesh.

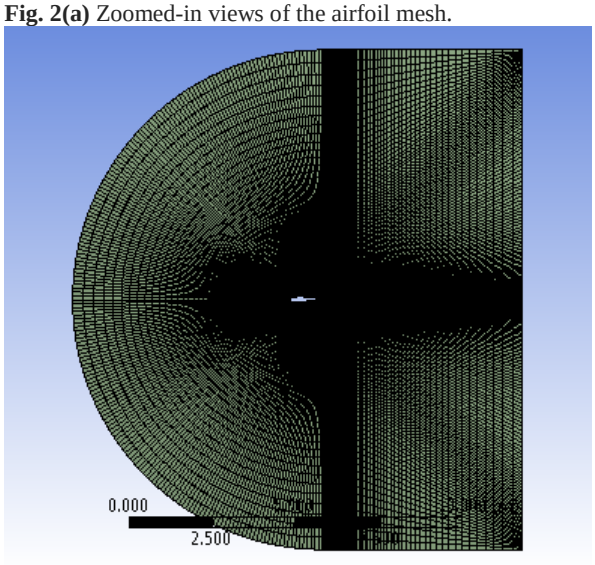


Fig. 2(b) Mesh of entire domain.

2.3. Inputs and Boundary Conditions

The problem considers flow around the airfoil at configurations of Angle of Attack (AOA) [0º, 4º, 8º, and 12º] and Re numbers [0.5 x 10^5, 1 x 10^5, and 2 x 10^5]. For that, some initial inputs and boundary condition are fed to the software and presented in the *Table 1*. Surface of airfoil is selected as a solid wall. The inlet boundary is given by using the free stream.

Table 1 Boundary conditions at the inlet.

S.no	Input	Value
1	Operating Temperature	288.16 Kelvin
2	Density of Fluid (air)	1.225 Kg/m3
3	Model	Spalart-Allmaras
4	Kinematic Viscosity	0.000177701
5	Length (Chord length)	1 m
6	Ratio of Specific Heats	1.4
7	Turbulent Intensity	0.3 %

2.4. Solver

All solutions in this study were computed using the Reynolds Averaged Navier-Stokes solution scheme. Numerical analysis was done using the density-based solver by applying the transport equations at the nodes of the grid elements. The inputs that have been specified to FLUENT for solving a problem are shown in *Table 2*. Convergence of solution varied in the different configurations, but the solution always converged in less than 2500 iterations in any of the cases.

Table 2. Inputs and Selections for this problem

S.no	Input	Selection
1	Type of solver	Density-based
2	Time	Steady
3	Formulation	Implicit
4	Gradient	Least Squares Based
5	Flow	Second Order Upwind
6	No. of Iterations	4000

3. Results

After successful calculations and results convergence, the graphs and contours were plotted and studied via *ANSYS CFD-POST*.

3.1 Grid Convergence

Grid convergence was checked via various grid sizes containing 10050 cells, 20100 cells, 40200 and 80400 cells respectively. Grid convergence was observed when the grid size consisted of 40200 cells.*Figure 3* represents a curve informing about the grid convergence in detail.

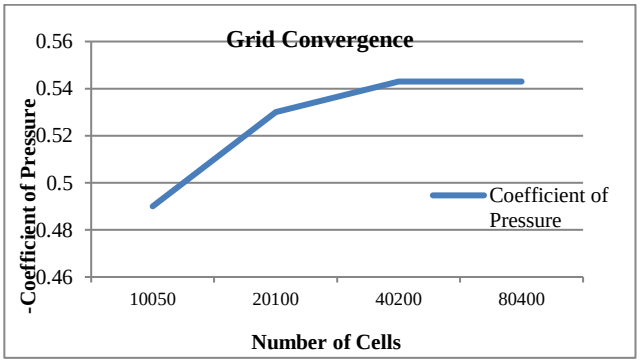


Fig. 3 Chart between values of –Cp obtained against the grid size. (Re = 2 x 105, Angle of Attack = 00 and Chord length = 0.2)

3.2 Pressure Contours

By the help of the Pressure contours from *Figures 4.a, 4.b* and *4.c*, it was observed that a region of high pressure was observed near the trailing edge on the lower surface and pressure gets reduced as we move away from the trailing edge along the surface. For zero angle of attack, there was less variation in pressure values at same chord lengths on the upper and lower surfaces. When the angle of attack is increased, more variation is observed in the respective values, lower surface pressures being considerably higher than the upper surface. As the Reynolds number is increased, pressure at all points also increases.

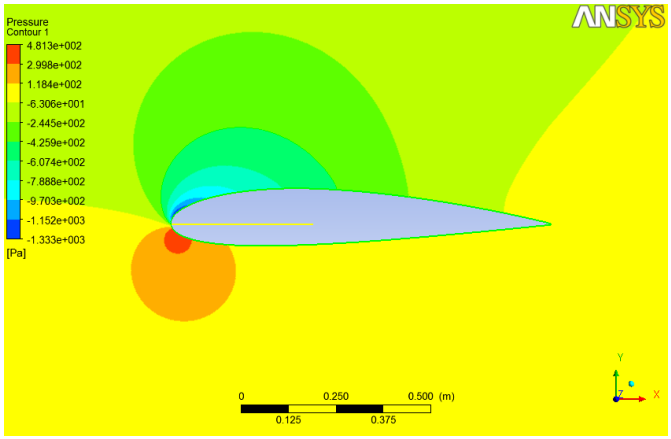


Fig. 4(a) Pressure contours across airfoil at Re=2 x $10^5$and Angle of Attack 12^0

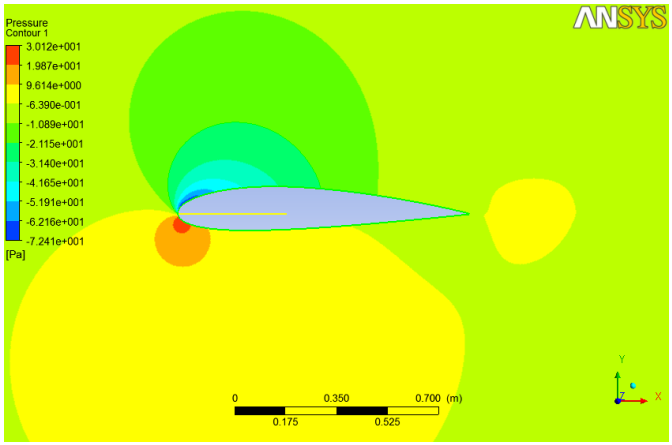


Fig. 4(b) Pressure contours across airfoil at Re = 0.5 x 10^5 and Angle of Attack 12^0

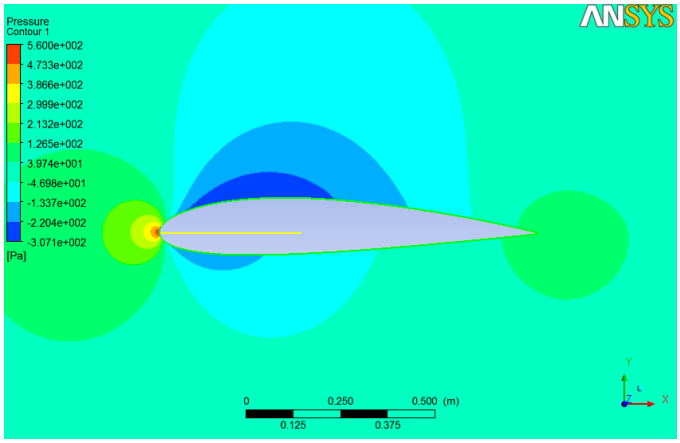


Fig. 4 (c) Pressure contours across airfoil at Re = 2 x 10^5 and Angle of Attack 0^0

3.3 Charts: Pressure v/s Chord length

Figures 5.a and *5.b* depict the definite value of pressure across the upper and lower surfaces of the airfoil. As said

above, when AOA was set to 0 degrees, there is significantly lesser difference between the pressures between upper and lower surfaces. At AOA = 12 degrees, pressure variation is observed to be maximum. Chord length of the airfoil is 1 unit (In this case, 1 metre).

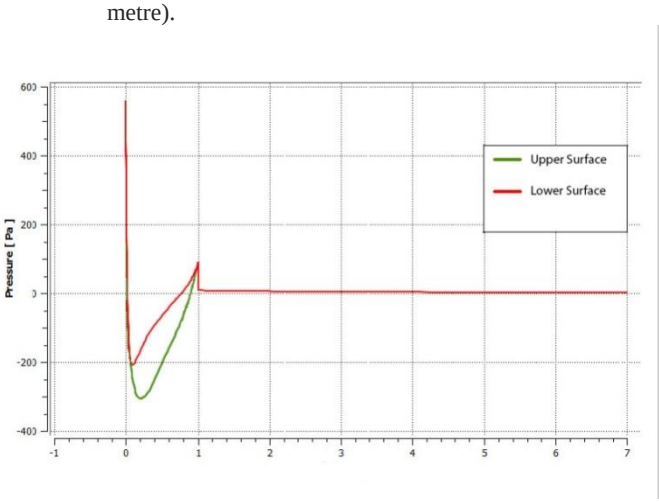


Fig. 5(a): Plot between Pressure and Chord length across Upper as well as Lower Surface of Airfoil (Re = 2 x 105, AOA = 00)

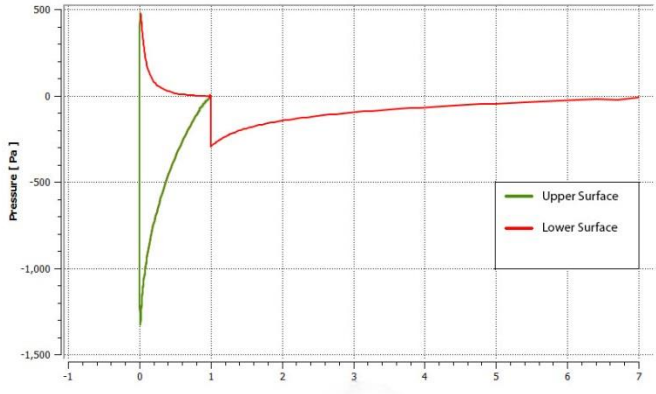


Fig.5(b) Plot between Pressure and Chord length across Upper as well as Lower Surface of Airfoil (Re = 2 x 105, AOA = 120.)

3.4 Velocity Contours

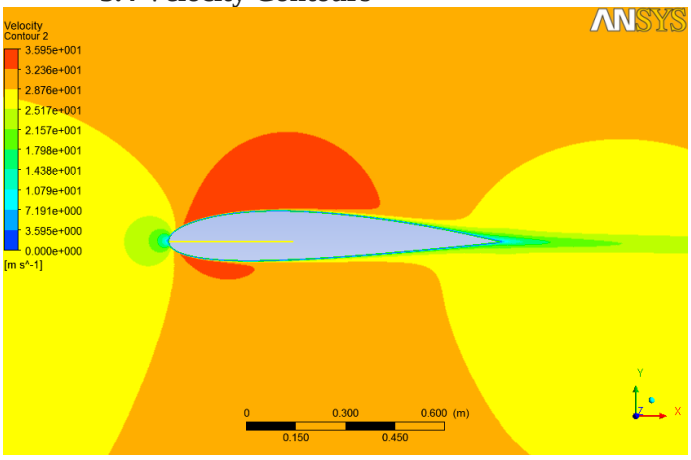


Fig. 6(a): Velocity contours across airfoil at Re = 2 x 10^5 and Angle of Attack 0^0

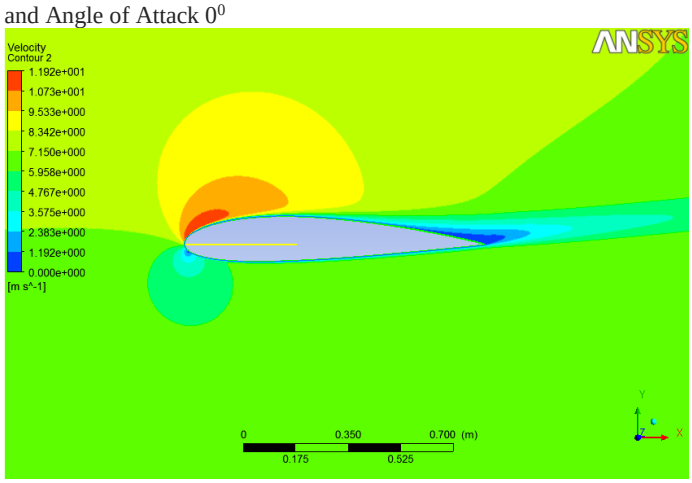


Fig. 6(b) Velocity contours across airfoil at Re = 0.5 x 10^5 and Angle of Attack 12^0

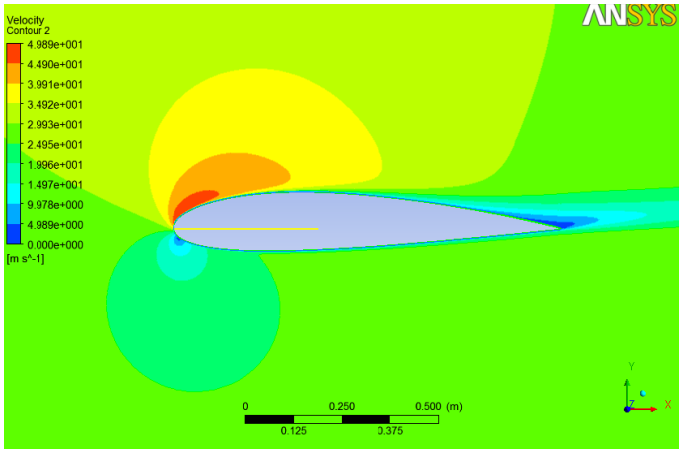


Fig. 6(c): Velocity contours across airfoil at Re = 2 x 10^5 and Angle of Attack 12^0

By the help of the velocity contours from Figures 6.a, 6.b and 6.c, it was observed that a region of high velocity was observed near the trailing edge on the upper surface and velocity gets reduced as we move away from the trailing edge along the surface. For zero angle of attack, there was less variation in velocity magnitudes at same chord lengths on the upper and lower surfaces. When the angle of attack is increased, more variation is observed in the respective values, lower surface pressures being considerably lower than the upper surface. As the Reynolds number is increased, velocity at all points also increases.

3.4 Comparison between Different Models

Figure 7. represents a chart where the experimental data, and CFD results obtained by both Spalart Alramus and K-Epsilon Methods are plotted together for observation and to draw out inferences.

It can be observed that the coefficient of pressure obtained when method Spalart Allmaras is chosen proves to be better and more accurate when compared against the method K-epsilon. Results obtained by Spalart Allmaras method show close resemblence to the values obtained by the experiement and also follow the trend of the experimental data curve.

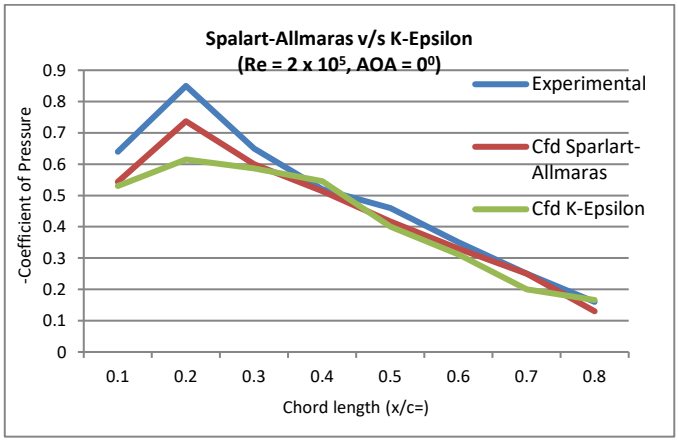
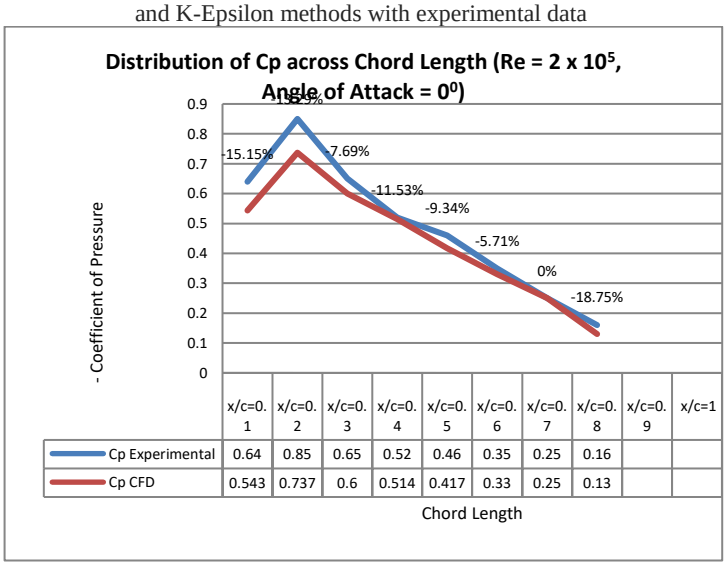
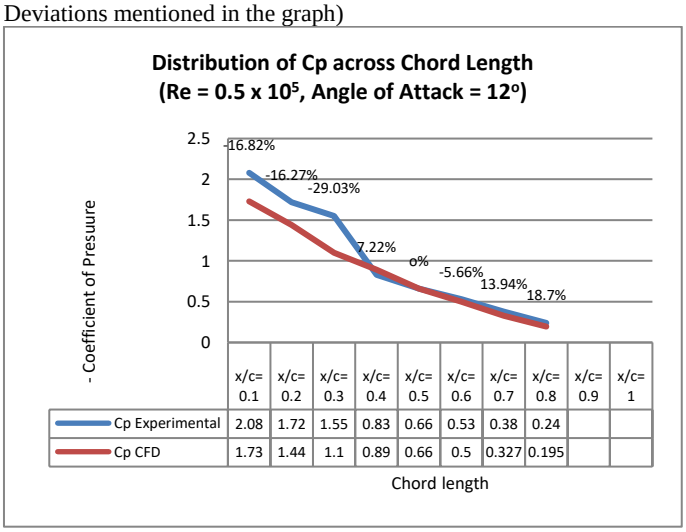


Fig.7 Chart for comparison of results of Spalart-Allmaras and K-Epsilon methods with experimental data



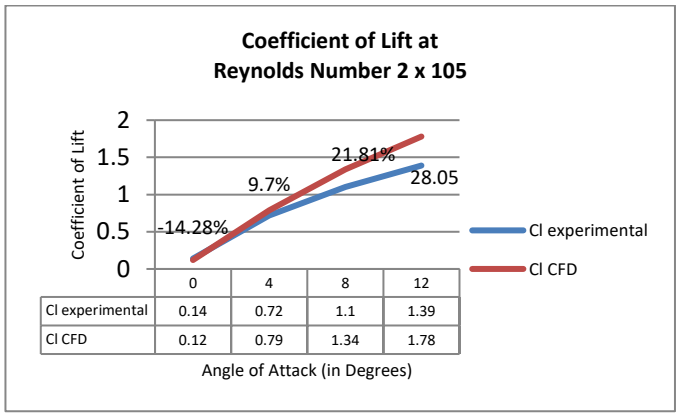
	x/c=0.1	x/c=0.2	x/c=0.3	x/c=0.4	x/c=0.5	x/c=0.6	x/c=0.7	x/c=0.8	x/c=0.9	x/c=1
Cp Experimental	0.64	0.85	0.65	0.52	0.46	0.35	0.25	0.16		
Cp CFD	0.543	0.737	0.6	0.514	0.417	0.33	0.25	0.13		

Fig. 8(a) Chart between –CP (Coefficient of Pressure) and Chord length for Reynolds number = 2 x 10^5 and Angle of attack = 00. Experimental Data as well as CFD results are plotted together (% Deviations mentioned in the graph)



	x/c=0.1	x/c=0.2	x/c=0.3	x/c=0.4	x/c=0.5	x/c=0.6	x/c=0.7	x/c=0.8	x/c=0.9	x/c=1
Cp Experimental	2.08	1.72	1.55	0.83	0.66	0.53	0.38	0.24		
Cp CFD	1.73	1.44	1.1	0.89	0.66	0.5	0.327	0.195		

Fig. 8(b) Chart between –CP (Coefficient of Pressure) and Chord length for Reynolds number = 0.5 x 105 and Angle of attack = 120. Experimental Data as well as CFD results are plotted together (% Deviations mentioned in the graph).



	0	4	8	12
Cl experimental	0.14	0.72	1.1	1.39
Cl CFD	0.12	0.79	1.34	1.78

Fig. 8(c) Chart between –CL (Coefficient of Lift) and Angle of Attack for Reynolds number = 2 x 105. Experimental Data as well as CFD results are plotted together (% Deviations mentioned in the graph)

3.5 Comparison of Results obtained with Experimental Data

The results obtained were compared to the experimental results of a research experiment conducted by University of Ecylies[9], on the same airfoil and similar conditions in a wind tunnel. The comparison as shown in Figures 8.a, 8.b and 8.c, showed close resemblence thus validating this research. Hence, the results of this experiment can be used for further investigations and experiments.

4.Conclusions

This project simulates NACA2415 airfoil on ANSYS FLUENT at low Reynolds numbers at different angles of attack. With Spalart-Allmaras as the preferred turbulent model, results were calculated and analyzed. Contours of Pressure, Velocity were presented and Coefficient of Pressure, Lift were compared against the experimental values. The following conclusions can be considered.

- Most appropriate turbulent model for this simulation found out to be Spalart-Allmaras Turbulent model when tested against other turbulent models such as k-epsilon. It yielded more accurate results.
- Contours of Pressure and Velocity show that parameters of Pressure and Velocity are greatly influenced by Reynolds number and Angle of Attack.
- Coefficient of Pressure (CP) and Coefficient of Lift (CL) were calculated and compared with the experimental values which showed close resemblance thereby, validating this research.

This study can further be optimized by testing it in unsteady states and on a wider range of angle of attack and Reynolds numbers.

References

[1.] SinghK.P, Mathur, JS, Ashok, V, Chakraborty D., Computational Fluid Dynamics in Aerospace Industry in India, Defence Science Journal, 60(6), 2010, 639-652.

[2.] Airfoil Nomenclature, Images, Wikipedia, http://en.wikipedia.org/wiki/Airfoil

[3.] 'Airfoil', Wikipedia, http://en.wikipedia.org/wiki/Airfoil\

[4.] Hasanuzzaman, Md. and Mashud, Md., Effect of Reynolds Number on Aerodynamic Characteristics of an Airfoil with Flap, Annals of Pure and Applied Mathematics, 3(1), 2013, 27-40.

[5.] Ghods M, Theory of wings and wind tunnel testing of a NACA 2415 airfoil, Technical Communication for Engineers, The University of British Columbia, 23, 2001.

[6.] Bin Ismail Md, Parametric investigation on the aerodynamic characteristics of NACA2415 airfoil at low Reynolds number, Faculty of Manufacturing Engineering, University of Malaysia, Pahang, 2013.

[7.] Velazquez-Araque1 L, Nožička J. Computational Analysis of the 2415-3S Airfoil Aerodynamic Performance, Systemics, Cybernetics and Informatics 12(1), 2014.

[8.] Turbulence Modeling Resource, Langley Research Center'; http://turbmodels.larc.nasa.gov/spalart.html

[9.] Serdar Genç, M., Karasu, I., Hakan Açıkel, H., 'An experimental study on aerodynamics of NACA2415 aerofoil at low Re numbers', M.S. Genç et al./ Experimental Thermal and Fluid Science 39, 2012, 252–264.

Computational Materials Science 130 (2017) 222–231

Contents lists available at ScienceDirect

Computational Materials Science

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

On the evolution and electronic properties of self-assembled gold nanowires

Neha Tyagi [a], Neeraj K. Jaiswal [b], Goran Kovačević [c,*], Pankaj Srivastava [d]

[a] Department of Applied Physics, Delhi Technological University, Delhi 110042, India
[b] Discipline of Physics, Indian Institute of Information Technology Design & Manufacturing, Dumna Airport Road, Jabalpur 482005, MP, India
[c] Rudjer Bošković Institute, Bijenička 54, P.O.B. 180, HR-10002 Zagreb, Croatia
[d] Nanomaterials Research Group, ABV-Indian Institute of Information Technology & Management, Gwalior 474015, India

ARTICLE INFO

Article history:
Received 20 July 2016
Received in revised form 30 November 2016
Accepted 7 January 2017

Keywords:
Gold nanowire
Self-assembly
Semi-metallic
Semiconducting
Metallic

ABSTRACT

The crucial importance for implementing nanomaterials to nanotechnology applications is their characterization in terms of structure and properties. Here we present electronic and conductive properties of seven different gold nanowires (1. zigzag, 2. rhomboid, 3. ladder, 4. c-ribbon, 5. h-ribbon, 6. pentagonal pyramid and 7. capped pentagonal bipyramid) that can be formed due to self assembly from the linear atomic chain. From density functional calculations backed up with molecular dynamics, nudged elastic band and geometry optimizations, it was found that, by reducing the tension on the linear atomic chain of gold atoms, it spontaneously self-assembles into one of the observed, energetically more stable motifs. It is revealed that as self assembly progresses, the electronic property of the motifs undergo semi-metallic → semiconducting → metallic transformations. The observed electronic behavior is further supported by the current-voltage (I–V) characteristics. It is noticed that few selective structures exhibit perfect linear behavior of I–V curve whereas rest follow typical semiconductor curve. Interestingly, pentagonal pyramid displays perfect negative differential resistance with a high peak to valley ratio (PVR) of 17.25 which ensures its potential for oscillators application. Our findings include the possibility of tailoring the electronic properties of gold NW during their evolution in self assembling process.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Nanowires (NW) are the typical 1-D structures which pledge to serve as building blocks in emerging nano-devices [1–4]. Their research is increasing continuously, covering wide areas, from basic properties to innovative applications [5–9]. It is known that the physical and electronic properties of NW are ruled by quantum-mechanical effects and are, therefore, highly dependent upon their diameter and shape [10]. Intensive research on NW has made it possible to synthesise them from different materials and on various substrates [11–13], however, obtaining a NW with desired electronic properties is yet a hard nut to crack. Owing to their unique 1-D confinement, attempts are being made to produce NW of controlled diameter, various shapes and surface orientations [11,14]. Also, efforts have been made to tailor their electronic properties via doping or adsorption of different atomic species [7,15,16]. Amongst all, metallic NW have attracted special attention of the research community due to potential for interconnections in upcoming nano-devices [17–20]. Their another advantage is that their diameter can be achieved up to mono atomic regime [21,18]. These mono atomic NW (also known as atomic chains) have been produced experimentally for a number of metals like Au, Co, Pt, Pd, etc. [18,19]. Coura et al. [19] reported the formation of linear atomic chains of gold resulting from stretching of gold NW. Hasmy et al. [22] studied the formation of suspended monoatomic chains of noble metals (Au, Ag, Cu) and revealed that Au is forming longer atomic chains than that of Ag and Cu. In an another study, Li chains exhibit potential to act as an atomic switch [23]. Fully spin polarized electrical transport was observed even at room temperature in the suspended chains of Co, Pd and Pt [18]. Besides these experimental reports, theoretical investigations have also been made to understand the evolution and potential applications of linear atomic chains. NW encapsulated in nanotube or absorbed on it has also been studied [24–26]. It is remarkable to note that previous studies are mainly focused on the linear atomic chains. On the other hand, the possibility of other deformed atomic chains are almost unexplored which can present interesting physics.

Motivated from this, in the present work, we have investigated the structural and electronic properties of Au atomic NW in various

* Corresponding author.
E-mail address: gkova@irb.hr (G. Kovačević).

http://dx.doi.org/10.1016/j.commatsci.2017.01.012
0927-0256/© 2017 Elsevier B.V. All rights reserved.

possible geometries evolving from the linear atomic chains. We simulated their self-assembly from the linear atomic chain and determined the mechanism and the energetics for each Au atomic NW in order to prove possibilities for their self-assembly. Our simulations of the self-assembly of a gold nanowire describe the reverse process, reported by Coura et al. [19]: instead of stretching of a NW to a linear atomic chain, we are letting a linear atomic chain to self-assemble by shrinking.

2. Computational method

Simulation of the self-assembly of gold nanowires was performed as a sequential series of geometry optimizations and molecular dynamics calculations. Geometry that was used in each step was created from the previous one in the sequence by contracting one dimension of the simulation cell by 5%. As an initial geometry, a single-atom thick gold nanowire (20 atoms long) was used with periodic boundary conditions applied. Tetragonal cell shape was used with the variable cell dimension parallel to the chain direction. Dimensions of the cell, perpendicular to the chain direction were fixed at 50 Å in order to prevent any significant interaction between chains from neighboring cells. General gradient approximation (GGA) in the form of the Perdew–Burke–Ernzerhof (PBE) [27] functional within the Density functional theory [28] was used in geometry optimization and molecular dynamics calculations. Siesta method, [29,30] implemented in the Siesta program package was used for calculations. Double ζ with polarization (DZP) basis set was used. The orbital cut-off radius was defined with the energy shift of 250 meV. Numeric evaluation of spatial properties, during DFT calculations were done on a grid whose size is determined by its plane-wave cutoff energy of 200 Ry. Geometry optimization was done with the conjugated gradient algorithm until forces were smaller than 0.03 eV/Å. MD calculations were done with Nose thermostat at constant temperature of 300 K and Nose mass of 50 Ry/fs, with time step of 1 fs. The simulated system with the size of 20 atoms is large enough to develop a great number of atomic configurations during the simulation procedure. That often created non-homogeneous systems, with more than one structural motif. The simulation of the self-assembly process can be seen as animation (Au_self_assembly.mp4) in supplementary material. In order to investigate the characteristics of self-assembled atomic motifs, the original system of 20 atoms was cut into smaller systems with less than 10 atoms, but with the more homogeneous atomic arrangement. Obtained systems were subjected to geometry optimization. For each optimized system, a trajectory with 11 successive geometries was created by the interpolation of atomic coordinates between the optimized Au nanowire motifs and the initial geometry of the same size (linear atomic chain), in order to gain insight into the mechanism of the self-assembly of gold nanowires. In-house implementation of the nudged elastic band (NEB) method [31] was used for the optimization of the reaction trajectories to the minimum energy path. The optimized initial trajectory with 11 geometries was used for creation of more detailed trajectory with 21 geometries. That, more detailed trajectory was created by inserting geometries between each neighboring pair of geometries in the initial trajectory with 11 geometries. Inserted geometries were created by interpolation of geometries of neighboring pairs. Siesta program was used for obtaining energies and energy gradients for all geometries during NEB optimizations.

For the computation of electronic structures, supercell models of self-assembled gold NW were used. The supercell was taken periodic in z-axis and confined with a vacuum region of 10 Å along x and y-axes. The Brillouin zone was sampled with 100 equispaced k point along the periodic direction. The transport properties were explored by incorporating density functional theory based nonequilibrium transport as implemented in Atomistix ToolKit Virtual Nanolab (ATK-VNL) [32]. ATK uses Landauer–Büttiker formalism $G = G_0 \sum_i T_i$ where G_0 $(= 2e^2/h)$ is the quantum of conductance and T_i is transmittance of the i^{th} channel at the Fermi level. First, the transmission spectra is calculated which accounts the total transmission across the Fermi level:

$$T(E, V) = T_r[\mathbf{t}^\dagger \mathbf{t}](\epsilon), \tag{1}$$

where $\mathbf{t}$ is the transmission amplitude matrix given by

$$\mathbf{t}(\epsilon) = [T_L(\epsilon)]^{1/2} \mathbf{G}(\epsilon) [T_R(\epsilon)]^{1/2}. \tag{2}$$

Once, transmission spectra has been calculated, the current is determined as:

$$I = G_0 \int_{-\infty}^{\infty} d\epsilon [n_F(\epsilon - \mu_L) - n_F(\epsilon - \mu_R)] T_r[\mathbf{t}^\dagger \mathbf{t}](\epsilon). \tag{3}$$

3. Results and discussion

3.1. Geometries

The self-assembly of gold nanowires is made under assumption of collapsing a nanowire in the form of the linear atomic chain of gold atoms (Fig. 1a). Although some of obtained geometries have the same number of atoms, they have different unit cell size (parallel to nanowires). Therefore these geometries correspond to nanowires with a different distensions. The first geometry that appears has the same bonding pattern as the linear atomic chain of gold atoms, but their atoms are arranged in the zigzag fashion (Fig. 1b and 01_zigzag.mp4 in supplementary materials). All atoms

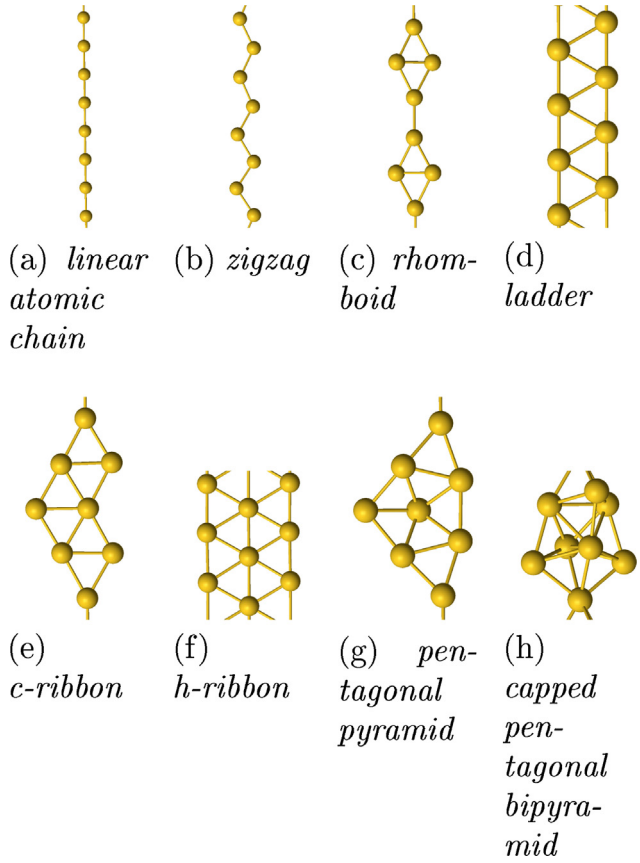


Fig. 1. Geometries of self-assembled gold nanowires.

in that geometry are organized around the same plane without significant deviations from the plane. Torsion angles defined from successive bonded atoms in that geometry do not exceed 6.5° from the perfect in-plane geometry. That geometry is the first geometry that the linear atomic chain of gold atoms gains upon reduction of tension. That makes it the precursor of all other geometries if the tension of the nanowire is further reduced. The simplest such rearrangement is the creation of rhomboid (Fig. 1c). The creation of the rhomboid from the zigzag conformation is accomplished with the creation of two new Au–Au bonds per rhomboid cluster. The entire process of self-assembly of rhomboid from the linear atomic chain, via the zigzag conformation, along the minimum energy path is optimized with the NEB method and geometries along its trajectory are shown as animation (02_rhomboid.mp4) in supplementary material. Rhomboid clusters are not collinear, but arranged in two plains that are aligned for about 44°.

If the unit cell dimension is reduced beyond the size convenient for rhomboid clusters, additional bonds are established between every other atom in a zigzag NW making the ladder conformation (Fig. 1d). The simplest way for obtaining the ladder conformation is the simple contraction of the zigzag NW. The process that happened in the simulation of the self-assembly of a gold nanowire (linear chain of 20 gold atoms) was an assembly by bending the Au chain around the rhomboid cluster and establishing bond with it. NEB optimization of the trajectory for the self-assembly of the ladder revealed yet another possibility for the creation of the ladder: by forming a triangle, followed by the zigzag collapsing of the gold chain (See animation 03_ladder.mp4 in supplementary material). That form of the self-assembly led to the rearrangement of atoms with the specific ordering, found in the final geometry of ladder, but different from the ordering resulting from the simple collapse of the zigzag. The observed self-assembly in the simulation of the c-ribbon (Fig. 1e) is a consequence of the collapse of the chain of gold atoms to the already assembled portion of the ladder. The mechanism is the same as the zigzag collapse of the linear chain to the ladder, except that the direction of bending of two atoms in the gold chain, in the initial stage of the collapse, is changed, disturbing the regular zigzag pattern and changing the direction of the growth of the ribbon. (See animation 04_c_ribbon.mp4 in supplementary materials.) The minimal energy path for the self-assembly of the c-ribbon follows the same pattern, except the portions of the NW are not assembled prior the final assembly into the c-ribbon, but the assembly is proceeding in concerted fashion. In the assembly, proceeding by the minimum energy path, the change in the zigzag pattern can be observed from the very beginning of the assembly process.

The c-ribbon is observed to rearrange (in the self-assembly simulation of a chain with 20 gold atoms) into the h-ribbon (Fig. 1f). The rearrangement is accomplished by braking of two Au–Au bonds and establishing three new Au–Au bonds. Atoms that contributed to the bond rearrangements established new bonds with neighboring atoms with respect to atoms, they lost bonding with. The minimum-energy path shows the same process of atom re-bonding in the last step for creation of h-ribbon, although the c-ribbon pattern was not observed in the same form in the NEB calculation. Animation of the minimum energy path for self-assembly of h-ribbon (05_h_ribbon.mp4) can be found in supplementary materials.

The formation of the pentagonal pyramid motif (Fig. 1g) also proceeds from the c-ribbon, but instead of bond rearrangements,

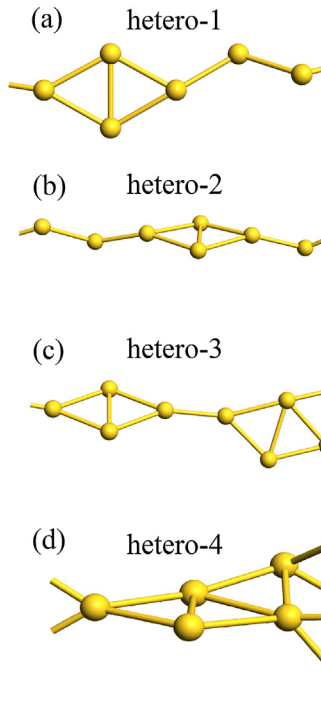


Fig. 2. The geometries of considered intermediate hetero-structures of gold nanowires.

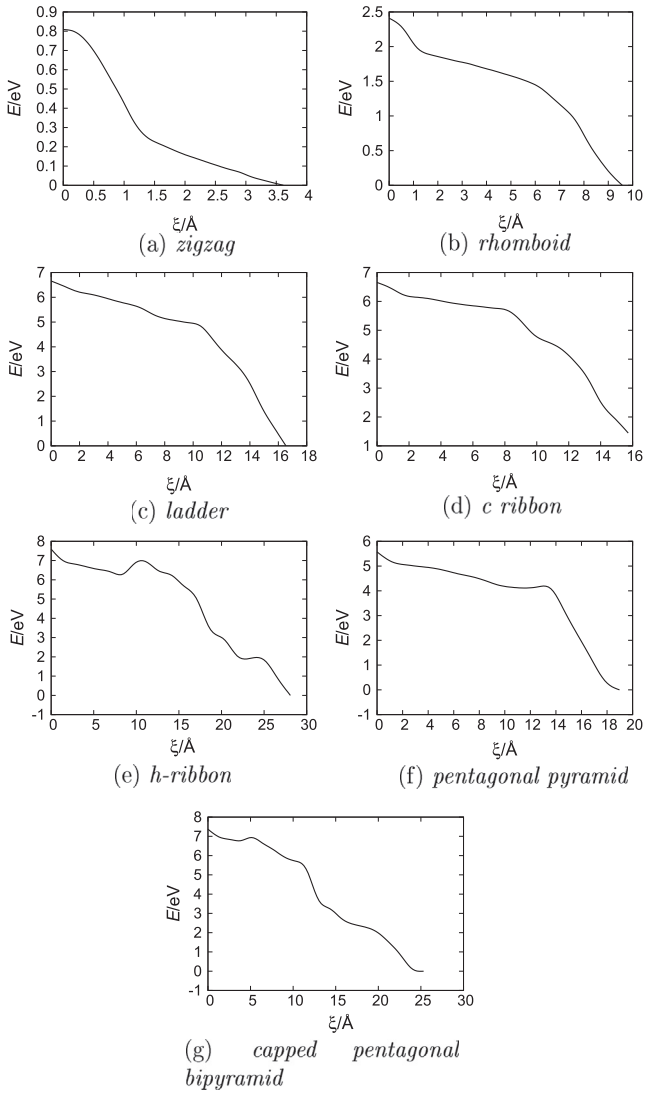


Fig. 3. Energy profiles for self-assembly of a linear atomic chain of gold atoms.

additional Au–Au bond is established. (See animation 06_pentagonal_pyramid.mp4 in supplementary materials.) For the establishment of the additional Au–Au bond, out-of-plain deformation of the c-ribbon is needed in order to bring two gold atoms close together for the formation of the bond. Rearrangement of the pentagonal pyramid motif leads to the capped pentagonal bipyramid (Fig. 1h). The minimum energy path trajectory show the formation of the c-ribbon as an intermediate, however it folds into the capped pentagonal bipyramid without formation of pentagonal pyramid. (See animation 07_capped_pentagonal_bipyramid.mp4 in supplementary materials.) In that case, the bond formation that characterizes the formation of the pentagonal pyramid also happens in the formation of the capped pentagonal bipyramid, but two other bonds form concertedly with it, so the bonding pattern of the c-ribbon converts toward the capped pentagonal bipyramid avoiding the pentagonal pyramid motif.

The stability of structural motifs can be corroborated by the similarity with isolated gold clusters [33–36]. All gold clusters with less than 7 atoms have most stable planar configurations [33,34]. The most stable gold cluster with four atoms is an isolated rhomboid and the most stable Au_5 cluster is a prototype for the ladder [33,34]. Pentagonal bipyramid can also be found as the most stable gold cluster with seven atoms [34]. In addition, gold clusters that are local minima in energy are equivalent for the pentagonal pyramid (Au_6) and the h-ribbon (Au_7) [33,34].

Since the proces of self-assembly of gold NW, guided by thermal motion of atoms does not necessarily proceed by minimum energy paths, non-homogeneous structures with more than one motif

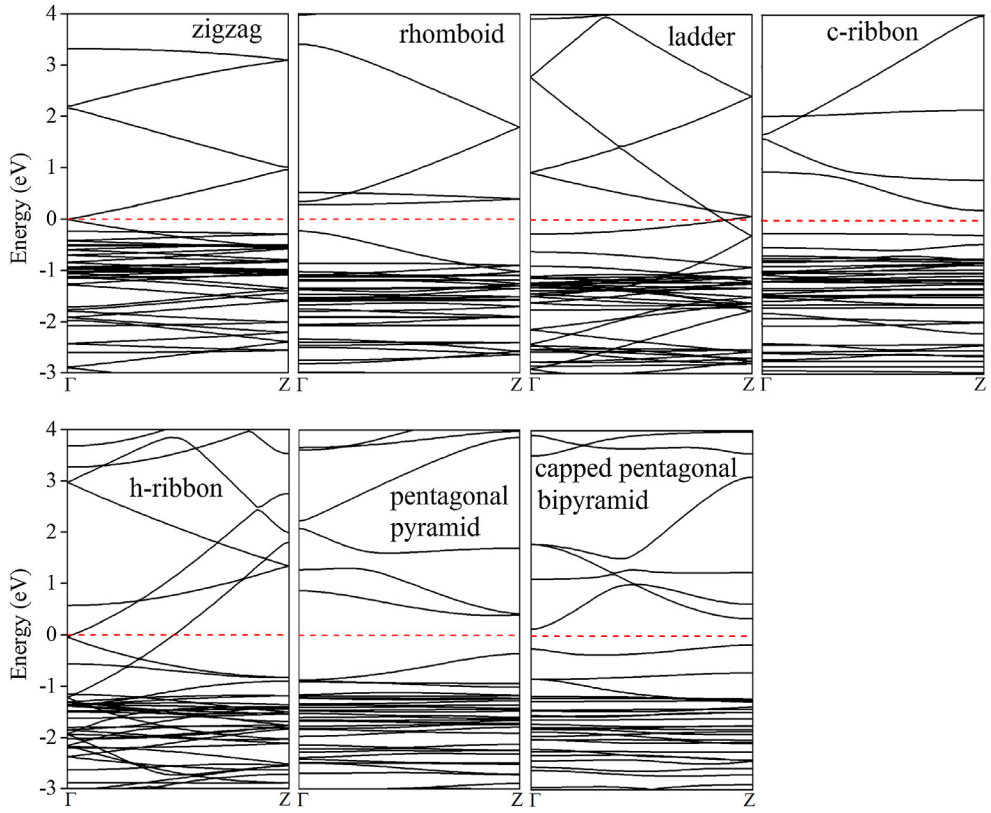


Fig. 4. The electronic band structures of considered gold NW as mentioned therein. The dashed line at 0 eV represents Fermi level.

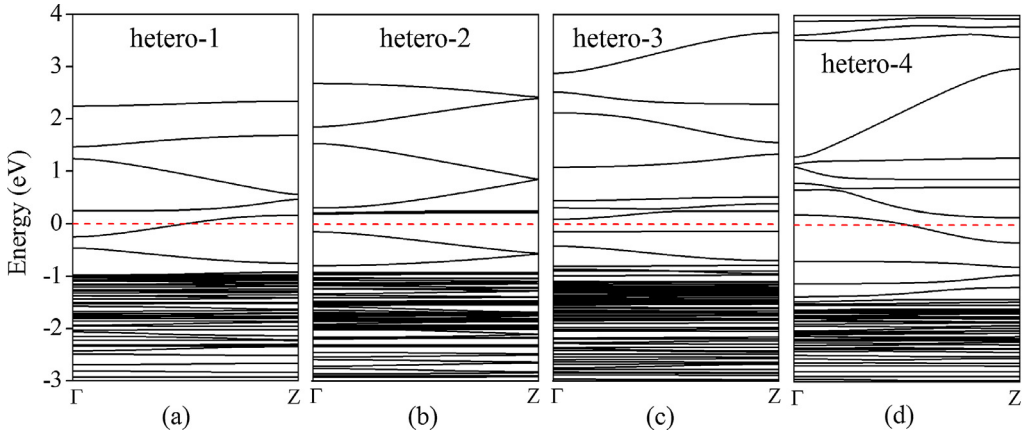


Fig. 5. The electronic band structures of intermediate hetero geometries of gold NW. The dashed line at 0 eV represents Fermi level.

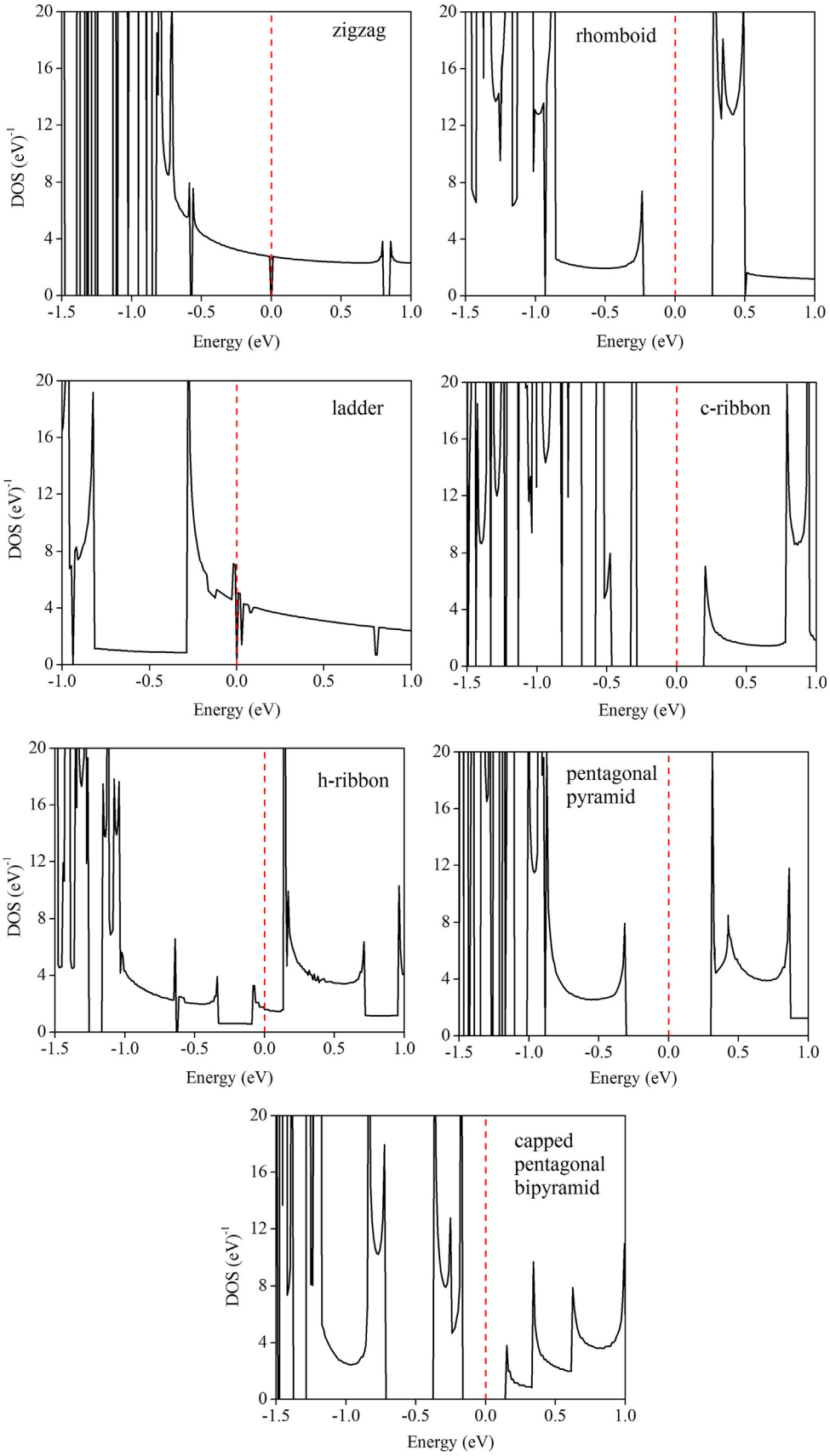


Fig. 6. The electronic density of states for considered gold NW as mentioned therein. The dashed line at 0 eV represents Fermi level.

present appear. In order to comment on the effect of possible defects, we consider the following heterostructures motifs evolved as an intermediate state during the self assembly of considered Au NW. These heterostructures are named as hetero-1, -2, -3, -4 as shown in Fig. 2. Their structures are taken from the trajectory of self-assembly of the gold nanowire (linear chain of 20 gold atoms) and submitted to geometry optimization and determination of their electronic structures, using the same numerical procedures as for homogeneous structures.

3.2. Energies

Optimization of minimum energy paths enabled the determination of energy profiles (Fig. 3) for self-assembly into each of the observed nanowire pattern. Reaction coordinates in trajectories (ξ) (Fig. 3) were calculated as a magnitude of the difference between atomic coordinate vectors in each geometry in the trajectory and the coordinate vector of the initial geometry (linear atomic chain). Thus, the reaction coordinate for j-th geometry in a trajectory (ξ_j) is:

$$\xi_j = \sqrt{\sum_{i=1}^{3N} (q_{i,j} - q_{0,j})^2}, \quad (4)$$

where N is the number of atoms in each geometry and $q_{i,j}$ is an i-th atomic coordinate in a j-th geometry of the trajectory.

Energy profiles in Fig. 3 correspond to minimum energy paths into nanowires in which each atom from the initial (linear atomic chain) conformation corresponds to the specific atom in the final conformation. There are possibilities of parallel reaction paths in which the initial arrangement of atoms result in exactly the same conformation of atoms in the resulting nanowire pattern, but with the different ordering of atoms (as discussed in previous section). Here we report reaction profiles that correspond to the minimum-energy path of the observed self-assembly from the simulation.

Each geometry in energy profiles in Fig. 3 has defined unit cell size, distinct to unit cell sizes of other geometries from the same trajectory. All unit cell sizes were frozen during NEB calculations. Unit cell sizes were chosen by linear variation between the unit cell sizes of the first and the last geometry in a trajectory. In all cases, geometries that result from self-assembly, upon the reduction of the unit cell size are significantly lower in energy than the initial conformation. That trend is expected since conformations that exist in shorter unit cells have more Au–Au bonds per atom than the initial conformation. Significantly lower energies of conformations with shorter unit cells make the self-assembly during shortening of nanowires, a spontaneous process. The simulated process of self-assembly is not completely free, since at each point there is a restriction put by the unit cell size. If that restriction would be removed, it is evident from energy profiles on Fig. 3 that each nanowire would spontaneously collapse towards

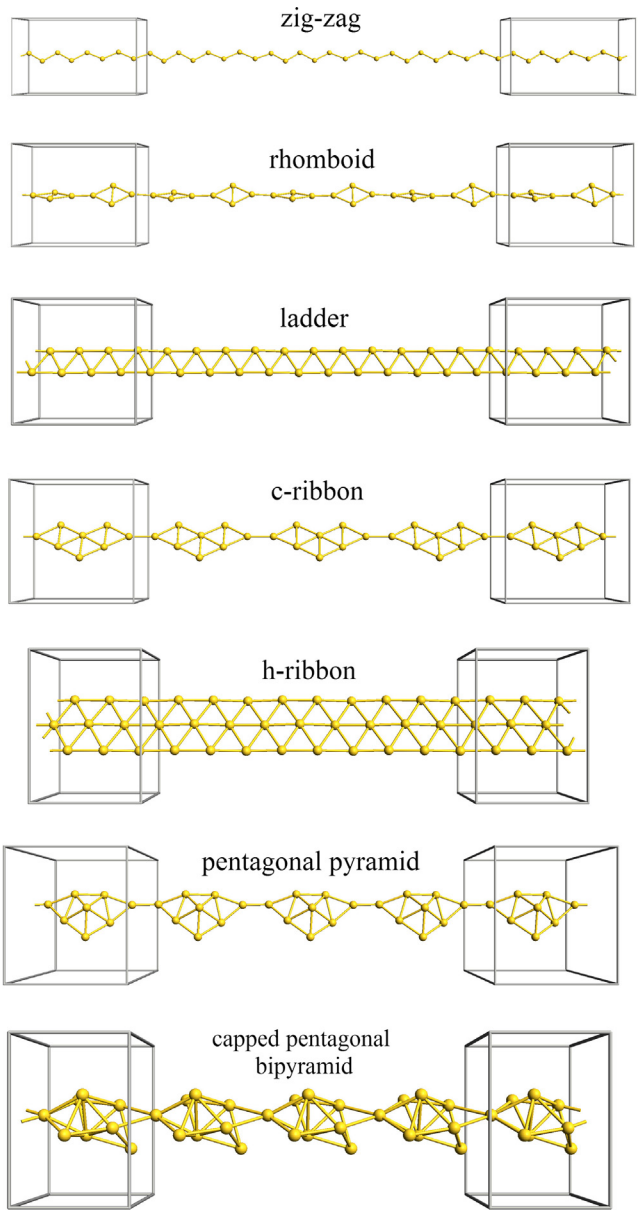


Fig. 8. The schematics of two-probe geometries used for the calculation of current-voltage (I–V) characteristics.

structures with lower energy and also smaller unit cells, resulting in clusters or perhaps a crystal structure. The restriction put by the unit cell size has the effect of the tension to the nanowire. Although energy stabilities of different structures of nanowires are significantly different, the tension (manifested as a defined unit

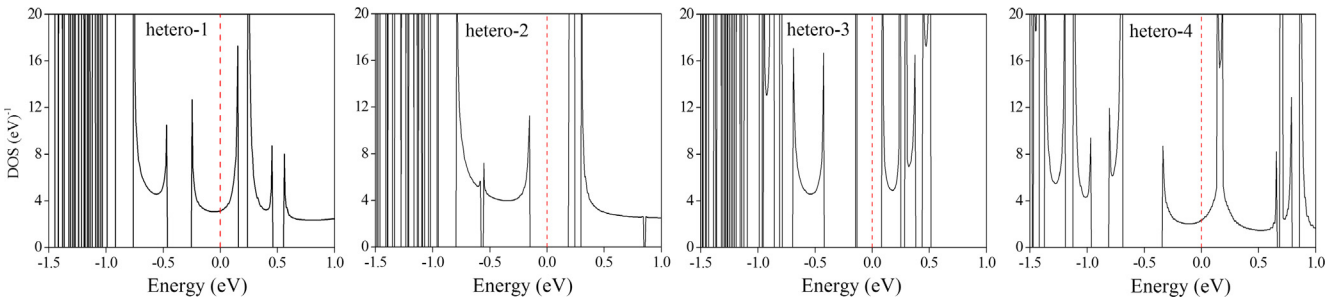


Fig. 7. The density of states profiles for intermediate hetero geometries of gold NW. The dashed line at 0 eV corresponds to Fermi level.

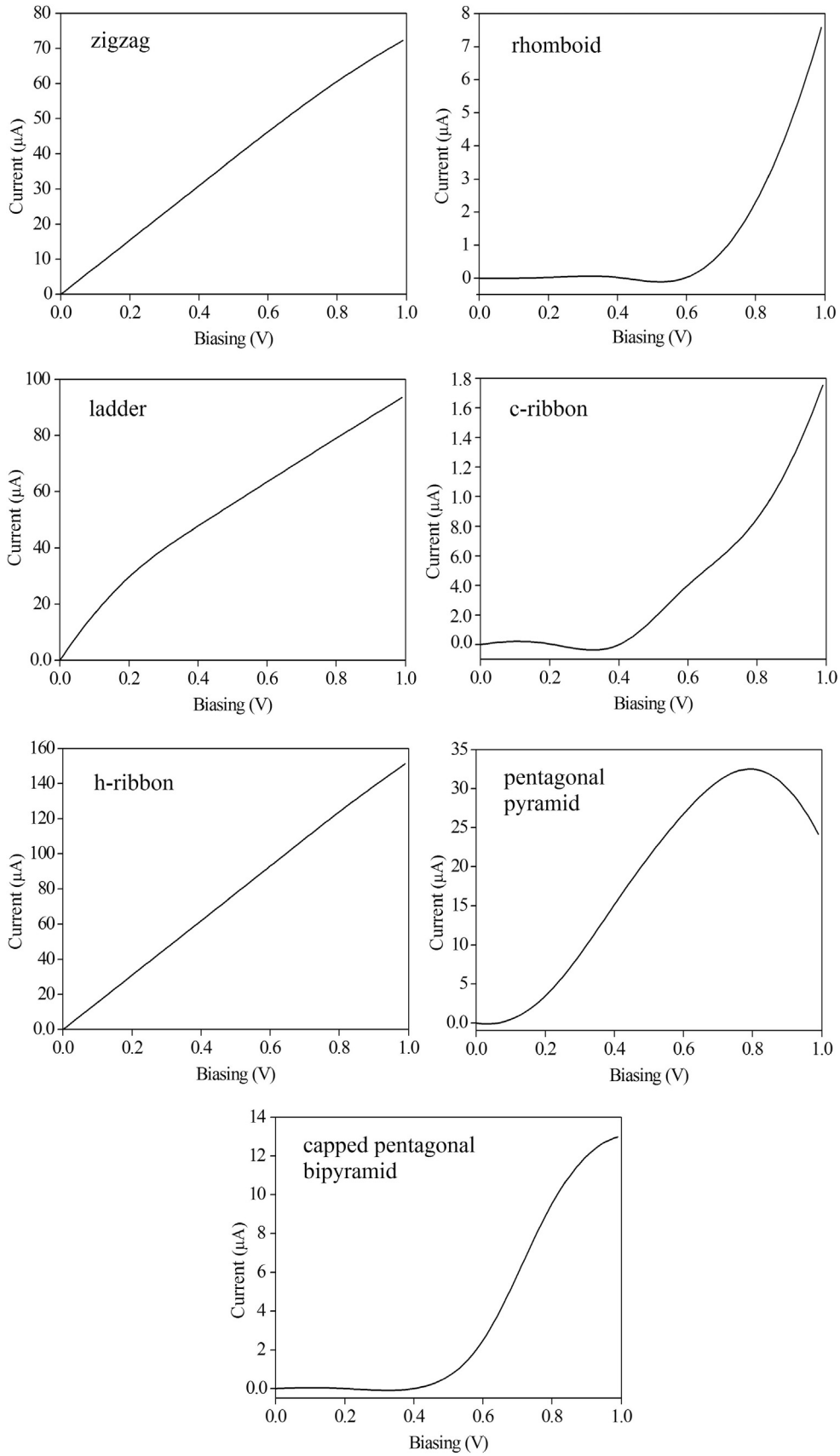


Fig. 9. The calculated current-voltage (I–V) characteristics for considered gold NW as mentioned therein.

cell size at the atomic level) applied to a nanowire disables conversions towards conformations with shorter unit cells and therefore enables the possibility for the control over a structure of the gold nanowire by applying a tension.

The observed mechanism of the self-assembly of rhomboid clusters conformation as the extension of the mechanism of self-assembly of the zigzag conformation is also evident in their energy profiles. The first 3.5 Å of the reaction coordinate for the self-assembly of the rhomboid clusters follows the same trend as the energy profile for the self-assembly of the zigzag conformation. The greatest gain in energy in that portion of the reaction path comes from the bending of gold atoms from the initial linear atomic chain. A significant energy stabilization in the formation of rhomboid clusters also comes from the formation of additional Au–Au bonds. All other reaction paths have similar initial portion since all start with bending of the initial linear atomic chain, similar to bending in zigzag fashion. The differences can be attributed to portions of nanowires whose folding departs from the ideal zigzag folding. In these reactions the portion that most significantly contributes to the gain in energy is the portion that takes place after the initial folding, when the majority of bonds are forming. Three of the trajectories (h-ribbon, pentagonal pyramid and capped pentagonal bipyramid) show local minima in energy profile during the self-assembly. These local minima are a consequence of bond elongation that cause rise in energy. In this cases, a non-linear change in the unit cell size (along the energy path) would prevent overstretched bonds and the temporary rise in energy during the self-assembly.

3.3. Electronic properties

To explore the electronic properties of considered Au NW, we study the electronic band structures as depicted in Fig. 4. It is revealed that in zigzag NW, the highest valence band (HVB) and lowest conduction band (LCB) just coincide at the Fermi level making it a zero band gap semi-metal. On the other hand, for diamond structure, we observed a direct band gap of 0.47 eV. The opening of band gap is attributed to the change in geometry from zigzag to rhomboid shaped. In case of ribbon, we observed the existence of electronic states across the Fermi level which makes this configuration purely metallic. Interestingly, c-ribbon, pentagonal pyramid and capped pentagonal bipyramid structures also exhibit the semiconducting nature and the magnitudes of band gap are found to be 0.50 eV, 0.75 eV and 0.41 eV respectively. The band structure of h-ribbon has two conducting channels across the Fermi level. The existence of these channels suggest enhanced metallicity for this configuration as compared to others.

The change in electronic properties for intermediate motifs (heterostructure NW) is depicted in Fig. 5. Interestingly, as the transition from zigzag to rhomboid is accomplished, the electronic properties undergo a semimetallic to metallic transition from zigzag to hetero-1 and then metallic to semiconducting transition from hetero-1 to hetero-2. Thereafter, the transition proceeds from hetero-2 structure to the rhomboid configuration. The redistribution of electronic states occurs during this transition and the band gap changes from 0.34 eV for hetero-2 to 0.47 eV in rhomboid NW. Next intermediate heterostructure (hetero-3) developed between rhomboid to ladder transition exhibits a band gap of 0.13 eV which lies in between that of 0.47 eV to 0 eV as observed for rhomboid and ladder NW respectively. On the other hand, pentagonal pyramid to capped pentagonal bipyramid structural transition is mediated via an intermediate metallic heterostructure (hetero-4).

For further verification of the observed electronic structures, we also examined electronic density of states (DOS) as shown in Fig. 6. The absence of electronic states across the Fermi level further supports the observed semiconducting behavior of rhomboid, c-ribbon, pentagonal pyramid and capped pentagonal bipyramid shaped NW. On the other hand, availability of DOS peaks in the vicinity of Fermi level is consistent with the metallic/semi-metallic character of zigzag, ladder and h-ribbon NW. Further, it is noticed that majority of density peaks for zigzag and capped pentagonal bipyramid configurations exhibit characteristic 0-D and 1-D behavior respectively. For rest of the NW structures, electronic peaks are admixture of 0-D and 1-D character.

The observed electronic structure predictions of heterostructures are further supported by their corresponding DOS profiles, depicted in Fig. 7. The absence of electronic states across the Fermi level for hetero-2 and heter-3 structures is consistent with their semiconducting nature as predicted by electronic band structures analysis. Further, availability of finite electronic states at Fermi level confirms the metallic character of hetero-1 and hetero-4 structures.

To examine the potential technological applications, we next investigate the current-voltage (I–V) characteristics of considered NW under finite biasing conditions. The schematics of the two-probe models used for I–V calculations have been presented in Fig. 8. In the geometry, the central region consists of three repetitions of the unit cell (as used for the band structure calculations) which is further coupled seamlessly with the electrodes. The calculated I–V curves have been depicted in Fig. 9. It is revealed that current is linearly dependent upon applied voltage for zigzag, ladder and h-ribbon NW. This linear dependency is due to metallic character of these NW as predicted from respective band structures [Fig. 4]. Moreover, the maximum current is noticed for h-ribbon NW since this configuration has larger number of conduction channels across the Fermi level. For rests of the NW, current remains almost zero until external biasing exceeds the respective band gaps which is in consistent with their semiconducting nature. Once the band gap has been overcame, current starts to increase with the applied voltage. However, for pentagonal pyramid NW, we observed a negative differential resistance (NDR) behavior [Fig. 9]. For this structure, current attains a maximum value of $\sim 32\ \mu A$ at about 0.8 V beyond which it starts to decline with increasing the biasing. Thus, a perfect NDR behavior is observed in pentagonal pyramid NW which make this configuration particularly useful for designing oscillators and fast switching devices. The maximum and minimum current values are noticed for h-ribbon and c-ribbon NW respectively. In order to use NDR phenomena for any device application, it is crucial to determine its peak to valley ration (PVR). Therefore, we plotted the I–V curve of pentagonal pyramid structure for an extended bias window of

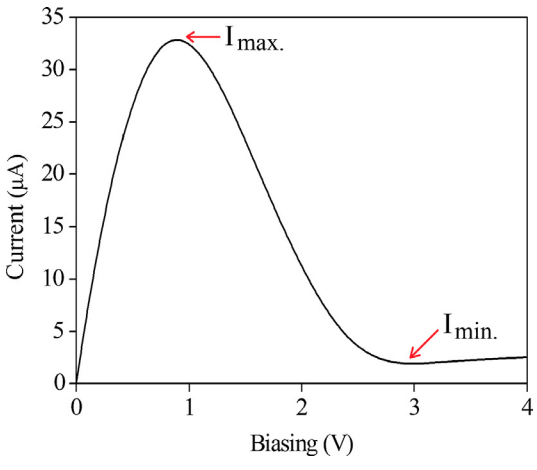


Fig. 10. The observed negative differential resistance (NDR) behavior for pentagonal pyramid NW.

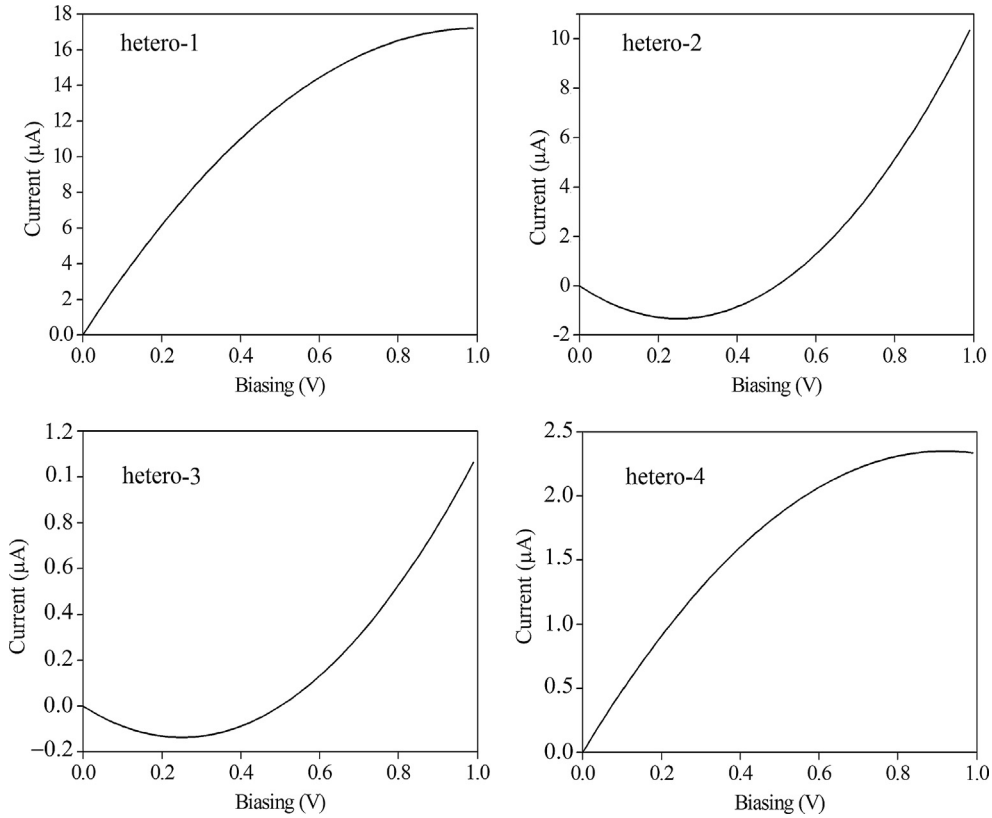


Fig. 11. The calculated I–V characteristics for intermediate hetero geometries of gold NW.

1–4 V as shown in Fig. 10. Our observed PVR for pentagonal pyramid NW is found to be 17.25 which is significantly higher than the previously reported PVR values for quantum-wire, interband tunnel diodes as well as 1-D derivatives of a new two-dimensional carbon allotrope [37–39]. This higher PVR value makes pentagonal pyramid NW a preferable candidate over previously studied structure for NDR device applications.

The analysis of I–V characteristics [Fig. 11] for heterostructures reveal typical semiconducting nature for hetero-2 and hetero-3 where the current starts to increase only after overcoming the corresponding band gaps of 0.34 eV and 0.13 eV respectively. For hetero-1 and hetero-4 structures, the current starts to increase without any delay, however, the maximum magnitude is noticed for hetero-1 motif.

The present calculated I–V characteristics exhibiting nonlinear behaviour is qualitatively consistent with that of experimental findings of Chandni et al. [40], however, the order of magnitude is different due to topological effects. Moreover, zigzag, ladder, h-ribbon, hetero-1 and hetero-4 exhibit almost linear character which is essentially in contrast with their finding of Tomonaga-Luttinger states.

4. Conclusion

In summary, we have examined the formation and electronic properties of self-assembled gold NW in different possible geometries. It is reported that various interesting motifs of gold can be derived out of its linear atomic string. Transformations of a linear atomic chain of gold atoms into observed NW is energetically favorable process that can be controlled by the external tension to the NW. The analysis of electronic properties revealed the possibility of semi-metallic → semiconducting → metallic transition depending upon the geometry of NW. The electronic structure of NW is also verified by the corresponding I–V characteristic. The zigzag, ladder and h-ribbon NW show linear dependency of current upon applied voltage. Moreover, the current for these structures is found significantly higher than that of remaining NW. On the other hand, for rhomboid, c-ribbon, and capped pentagonal bipyramid NW current remains quite low due to their semiconducting nature. Interestingly, pentagonal pyramid NW exhibit perfect NDR behavior with a PVR of 17.25. The observed NDR and linear I–V behavior ensure the potential of proposed gold NW respectively for oscillators and nano interconnect applications.

5. Contributions

NT did electronic structure calculations and wrote the manuscript. NKJ calculated I–V characteristic and analysed the results. GK wrote part of the manuscript, did simulations of the self-assembly of gold nanowires, MD and NEB calculations. PS wrote the part of manuscript.

Acknowledgments

NKJ is thankful to Indian Institute of Information Technology, Design and Manufacturing Jabalpur for providing infrastructural facilities and ABV-IIITM Gwalior for computational resources. This work has been partially suppotred by the Croatian Science Foundation under the project No. 6135.

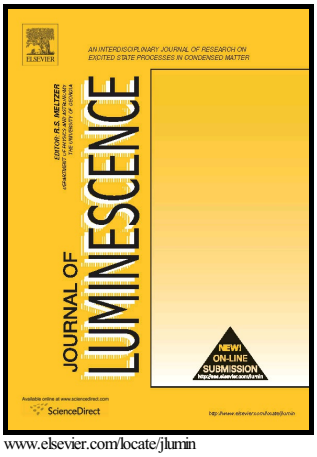
Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.commatsci.2017.01.012.

References

[1] Y. Cui, C.M. Lieber, Functional nanoscale electronic devices assembled using silicon nanowire building blocks, Science 291 (2001) 851–853.
[2] X. Duan, Y. Huang, Y. Cui, J. Wang, C.M. Lieber, Indium phosphide nanowires as building blocks for nanoscale electronic and optoelectronic devices, Nature 409 (2001) 66–69.
[3] S. Jeong, M. Choe, J.-W. Kang, M.W. Kim, W.G. Jung, Y.-C. Leem, J. Chun, B.-J. Kim, S.-J. Park, High-performance photoconductivity and electrical transport of ZnO/ZnS core/shell nanowires for multifunctional nanodevice applications, Appl. Mater. Interfaces 6 (2014) 6170–6176.
[4] Q. Su, Y. Zhang, S. Li, L. Du, H. Zhao, X. Liu, X. Li, Synthesis of p-type phosphorus doped ZnSe nanowires and their applications in nanodevices, Mater. Lett. 139 (2015) 487–490.
[5] J.C. Johnson, H.-J. Choi, K.P. Knutsen, R.D. Schaller, P. Yang, R.J. Saykally, Single gallium nitride nanowire lasers, Nat. Mater. 1 (2002) 106–110.
[6] J.T. Arantes, A.J.R. da Silva, A. Fazzio, Structural, electronic and magnetic properties of Mn-doped Ge nanowires by ab initio calculations, Phys. Rev. B 75 (2007) 115113-1–115113-6.
[7] C.M. Lieber, Z.L. Wang, Functional nanowires, MRS Bull. 32 (2007) 99–104.
[8] T.M. Schmidt, Surface effects on the energetic and spintronic properties of InP nanowires diluted with Mn: first-principles calculations, Phys. Rev. B 77 (2008) 085325-1–085325-6.
[9] E. Gadret, G.O. Dias, L.C.O. Dacal, M.M. de Lima, C.V.R.S. Ruffo, F. Iikawa, M.J.S.P. Brasil, T. Chiaramonte, M.A. Cotta, L.H.G. Tizei, D. Ugarte, A. Cantarero, Valence-band splitting energies in wurtzite InP nanowires: photoluminescence spectroscopy and ab initio calculations, Phys. Rev. B 82 (2010) 125327-1–125327-5.
[10] D.J. Carter, J.D. Gale, B. Delley, C. Stampfl, Geometry and diameter dependence of the electronic and physical properties of GaN nanowires from first principles, Phys. Rev. B 77 (2008) 115349-1–115349-12.
[11] R.C. Wang, C.P. Liu, J.L. Huang, S.-J. Chen, ZnO hexagonal arrays of nanowires grown on nanorods, Appl. Phys. Lett. 86 (2005) 251104-1–251104-3.
[12] M. Heurlin, P. Wickert, S. Falt, M. Borgstro, K. Deppert, L. Samuelson, M. Magnusson, Axial InP nanowire tandem junction grown on a silicon substrate, Nano Lett. 11 (2011) 2028–2031.
[13] J. Benson, S. Boukhalfa, A. Magasinski, A. Kvit, G. Yushin, Chemical vapor deposition of aluminum nanowires on metal substrates for electrical energy storage applications, ACS Nano 6 (2012) 118–125.
[14] J. Mallet, M. Molinari, F. Martineau, F. Delavoie, P. Fricoteaux, M. Troyon, Growth of silicon nanowires of controlled diameters by electrodeposition in ionic liquid at room temperature, Nano Lett. 8 (2008) 3468–3474.
[15] E. Durgun, N. Akman, C. Ataca, S. Ciraci, Atomic and electronic structures of doped silicon nanowires: a first-principles study, Phys. Rev. B 76 (2007) 245323-1–245323-8.
[16] P. Srivastava, A. Kumar, N.K. Jaiswal, Structural, electronic and magnetic properties of Mn-Doped InP nanowire, Superlatt. Microstruct. 92 (2016) 134–142.
[17] A.I. Yanson, G.R. Bollinger, H.E. van den Brom, N. Agrait, J.M. van Ruitenbeek, Formation and manipulation of a metallic wire of single gold atoms, Nature 395 (1998) 783–785.
[18] V. Rodrigues, J. Bettini, P.C. Silva, D. Ugarte, Evidence of spontaneous spin-polarized transport in magnetic nanowires, Phys. Rev. Lett. 91 (2003) 096801-1–096801-4.
[19] P.Z. Coura, S.B. Legoas, S. Moreira, A.F. Sato, V. Rodrigues, S.O. Dantas, D. Ugarte, D.S. Galvão, On the structural and stability features of linear atomic suspended chains formed from gold nanowires stretching, Nano Lett. 4 (2004) 1187–1191.
[20] Y. Zhou, S. Sreekala, P.M. Ajayan, S.K. Nayak, Resistance of copper nanowires and comparison with carbon nanotube bundles for interconnect applications using first principles calculations, J. Phys. Cond. Matter 20 (2008) 095209-1–095209-5.
[21] V. Rodrigues, D. Ugarte, Real-time imaging of atomistic process in one-atom-thick metal junctions, Phys. Rev. B 63 (2001) 073405-1–073405-4, http://dx.doi.org/10.1103/PhysRevB.63.073405.
[22] A. Hasmy, L. Rincón, R. Hernández, V. Mujica, M. Márquez, C. González, On the formation of suspended noble-metal monatomic chains, Phys. Rev. B 78 (2008) 115409-1–115409-5.
[23] T. Wunderlich, B. Akgenc, U. Eckern, C. Schuster, U. Schwingenschlögl, Modified Li chains as atomic switches, Sci. Rep. 3 (2013) 1–6.
[24] W.Y. Choi, J.W. Kang, H.J. Hwang, Structures of ultrathin copper nanowires encapsulated in carbon nanotubes, Phys. Rev. B 68 (2003) 193405-1–193405-4.
[25] E. Durgun, S. Ciraci, Spin-dependent electronic structure of transition-metal atomic chains adsorbed on single-wall carbon nanotubes, Phys. Rev. B 74 (2006) 125404-1–125404-8.
[26] L.-C. Ma, H.-S. Zhao, W.-J. Yan, Structural, electronic and magnetic properties of linear monoatomic chains adsorption on beryllium oxide nanotube: first-principle study, J. Magn. Magn. Mater. 330 (2013) 174–180.
[27] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865–3868, http://dx.doi.org/10.1103/PhysRevLett.77.3865.
[28] R.O. Jones, O. Gunnarsson, The density functional formalism, its applications and prospects, Rev. Mod. Phys. 61 (1989) 689–746, http://dx.doi.org/10.1103/RevModPhys.61.689.
[29] P. Ordejón, E. Artacho, J.M. Soler, Self-consistent order-n density-functional calculations for very large systems, Phys. Rev. B 53 (1996) R10441–R10444, http://dx.doi.org/10.1103/PhysRevB.53.R10441.
[30] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, D. Sánchez-Portal, The siesta method for ab initio order-n materials simulation, J. Phys.: Cond. Matter 14 (11) (2002) 2745–2779.
[31] G. Henkelman, H. Jónsson, Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points, J. Chem. Phys. 113 (22) (2000) 9978–9985, http://dx.doi.org/10.1063/1.1323224.
[32] M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, K. Stokbro, Density-functional method for nonequilibrium electron transport, Phys. Rev. B 65 (2002) 165401-1–165401-17.
[33] G. Bravo-Pérez, I. Garzón, O. Novaro, Ab initio study of small gold clusters, J. Molec. Struct.: {THEOCHEM} 493 (1–3) (1999) 225–231, http://dx.doi.org/10.1016/S0166-1280(99)00243-2.
[34] J. Wang, G. Wang, J. Zhao, Density-functional study of $au_n (n = 2 - 20)$ clusters: lowest-energy structures and electronic properties, Phys. Rev. B 66 (2002) 035418-1–035418-6, http://dx.doi.org/10.1103/PhysRevB.66.035418.
[35] H. Häkkinen, U. Landman, Gold clusters (au_N, $2 <; n <; 10$) and their anions, Phys. Rev. B 62 (2000) R2287–R2290, http://dx.doi.org/10.1103/PhysRevB.62.R2287.
[36] H. Grönbeck, W. Andreoni, Gold and platinum microclusters and their anions: comparison of structural and electronic properties, Chem. Phys. 262 (1) (2000) 1–14, http://dx.doi.org/10.1016/S0301-0104(00)00294-9.
[37] T. Sugaya, K.-Y. Jang, C.-K. Hahn, M. Ogura, K. Komori, A. Shinoda, K. Yonei, Enhanced peak-to-valley current ratio in *InGaAs*/*InAlAs* trench-type quantum-wire negative differential resistance field-effect transistors, J. Appl. Phys. 97 (2005) 034507-1–034507-5.
[38] A. Ramesh, P.R. Berger, R. Loo, High 5.2 peak-to-valley current ratio in Si/SiGe resonant interband tunnel diodes grown by chemical vapor deposition, Appl. Phys. Lett. 100 (2012) 092104-1–092104-4.
[39] B. Mandal, S. Sarkar, A. Pramanik, P. Sarkar, Theoretical prediction of a new two-dimensional carbon allotrope and NDR behaviour of its one-dimensional derivatives, Phys. Chem. Chem. Phys. 15 (2013) 21001–21006.
[40] U. Chandni, P. Kundu, A.K. Singh, N. Ravishankar, A. Ghosh, Insulating state and breakdown of fermi liquid description in molecular-scale single-crystalline wires of gold, ACS Nano 5 (10) (2011) 8398–8403, http://dx.doi.org/10.1021/nn2031935, pMID: 21942565.

Author's Accepted Manuscript



www.elsevier.com/locate/jlumin

Optimization of Synthesis Technique and Luminescent Properties in Eu^{3+}-activated $NaCaPO_4$ Phosphor for Solid State Lighting Applications

B.V. Ratnam, Mukesh K. Sahu, Amit K. Vishwakarma, Kaushal Jha, Hyun-Joo Woo, Kiwan Jang, M. Jayasimhadri

PII: S0022-2313(16)30987-5
DOI: http://dx.doi.org/10.1016/j.jlumin.2017.01.006
Reference: LUMIN14475

To appear in: *Journal of Luminescence*

Received date: 26 July 2016
Revised date: 12 December 2016
Accepted date: 5 January 2017

Cite this article as: B.V. Ratnam, Mukesh K. Sahu, Amit K. Vishwakarma, Kaushal Jha, Hyun-Joo Woo, Kiwan Jang and M. Jayasimhadri, Optimization o Synthesis Technique and Luminescent Properties in Eu^{3+}-activated NaCaPO Phosphor for Solid State Lighting Applications, *Journal of Luminescence* http://dx.doi.org/10.1016/j.jlumin.2017.01.006

This is a PDF file of an unedited manuscript that has been accepted fo publication. As a service to our customers we are providing this early version o the manuscript. The manuscript will undergo copyediting, typesetting, an review of the resulting galley proof before it is published in its final citable form Please note that during the production process errors may be discovered whic could affect the content, and all legal disclaimers that apply to the journal pertain

ACCEPTED MANUSCRIPT

Optimization of Synthesis Technique and Luminescent Properties in Eu^{3+}-activated $NaCaPO_4$ Phosphor for Solid State Lighting Applications

B.V. Ratnam[1], Mukesh K. Sahu[2], Amit K. Vishwakarma[2], Kaushal Jha[2], Hyun-Joo Woo[1], Kiwan Jang[1,*], M. Jayasimhadri[2,]**

[1]Department of Physics, Changwon National University, Changwon- 51140, Republic of Korea

[2]Luminescence Materials Research Lab, Department of Applied Physics, Delhi Technological University, Delhi-110042, India

kwjang@changwon.ac.kr (Kiwan Jang)

jayaphysics@yahoo.com (M. Jayasimhadri)

***Corresponding author.** Tel: +82-55-213-3427.

****Corresponding author. T**el.: +91- 9013553360

Accepted manuscript

Abstract

Europium activated $NaCaPO_4$ phosphor has been synthesized by various synthesis techniques such as solid-state reaction (SSR), molten salt synthesis (MSS) and sol-gel combustion (SGC) method to optimize the synthesis procedure. The comparative investigations of structural and luminescent properties have been studied to know the best synthesis method. The XRD patterns and Rietveld refinement analysis of the synthesized phosphors confirmed the single phase orthorhombic structure of $NaCaPO_4$. Excitation spectra indicate the strong absorption in near ultraviolet (n-UV) region and the emission spectra exhibit strong emission band at 595 nm corresponds to $^5D_0 \rightarrow {}^7F_1$ transition under n-UV (λ_{ex}=392 nm) excitation. The SGC route synthesized phosphor exhibit intense emission than that of the SSR and MSS method. Therefore, the effect of dopant (Eu^{3+}) concentration on the emission intensity and concentration quenching mechanism has been discussed in detail for the Eu^{3+} doped $NaCaPO_4$ phosphor synthesized by SGC method. The CIE chromaticity coordinates have been calculated for the phosphors synthesized by SSR, MSS and SGC

ACCEPTED MANUSCRIPT

methods to reveal the emitting color and also to know the utility of this phosphor for white LEDs.

Keywords: $NaCaPO_4$; Phosphor; Rietveld refinement; Luminescence; white LEDs

1. Introduction

Recently, there is an increasing demand for phosphor converted (pc) white light emitting diodes (LEDs) in the field of solid state lighting due to their excellent advantages such as small size, high energy efficiency, energy-saving, robustness, high brightness, fast switching, longer lifetime (>1,00,000 h) and environmental friendliness[1–3]. The first commercial method to obtain the phosphor converted white light emitting diode (pc-WLED) involves the combination of blue LEDs with yellow emitting cerium-doped yttrium aluminum garnet (YAG: Ce^{3+}) phosphor. However, this technique is suffering with some disadvantages like hallo effect due to the separation of blue-yellow color, thermal quenching and poor color rendering index (R_a) [4–6]. The second alternative method to generate better white light is the composition of the three RGB (i.e. red, green and blue) phosphors with a n-UV LED chip as an excitation source. However, in this approach, the efficiency of red phosphor is very poor as compare to the green and blue [7,8]. Therefore, it is necessary to investigate chemically stable, environmental friendly and highly efficient n-UV or blue excited phosphors to use in pc-converted WLED [9,10].

Inorganic alkali-alkaline-earth orthophosphate with general formula $ABPO_4$ (where A= alkali and B= alkaline earth metals) have been studied widely as efficient host materials due to their excellent luminescent properties [11–14]. In orthophosphate group, $NaCaPO_4$ is known as emerging host lattice for potential application in solid-state lighting [15,16]. Along with excellent thermal and charge stability, $NaCaPO_4$ also possess some unique qualities

Accepted manuscript

ACCEPTED MANUSCRIPT

such as strong UV absorption, high conversation efficiency with low cost, lower synthesis temperature, higher physical and chemical stability[17–19]. The doping of rare-earth ions in suitable orthophosphate host lattices have been considered as an excellent material for luminescent device applications [20]. Therefore, incorporation of rare-earth ions provide strong emission intensity as well as stability of the oxide materials. Among all the rare-earth elements Eu^{3+} ion is most commonly used activator in many host lattices for the desired red luminescence properties. Generally, Eu^{3+} ion can enter in the host lattice either in its divalent states or as trivalent states [20,21].

Various synthesis methods such as SSR, MSS, SGC, hydrothermal, combustion and solvothermal have been used to synthesize the phosphor materials. Among all syntehses, SSR method is most popular and commercially adopted technique to obtain the efficient phosphors. High firing temperature for a long duration, increase the crystallinity of phosphor, but same time degrade the morphological and structural properties [22,23]. Either molten salt or flux are used to provides a better medium for diffusion in solid state reaction, which may enhance the property of the product, and this method is generally known as molten salt or eutectic method [24,25]. In the molten salt method, the alkali metals are added in precursor mixture during the synthesis process. The alkali metal provides favorable reaction medium, it may enhance the kinetics of formation of the desired material. Therefore, the addition of a molten salt during the synthesis process can promote particle growth and improve the morphology of the particle with eliminating surface defects, which in turn enhances the PL emission intensity [26]. Nowadays, sol-gel combustion technique is very emerging and attractive method to synthesize the materials in nano or deep micron size [27,28]. The main advantage of this method is low cost, less time consuming, easy synthesis and high yield of product after complete reaction. However, the above mentioned synthesis routes have been used in the present work [29,30]. Various rare earth doped orthophosphates

Accepted manuscript

ACCEPTED MANUSCRIPT

$NaZnPO_4:Mn^{2+}$ [11,31], $KSrPO_4$ [32], $Ca_3(PO_4)_2$ [33] $NaMg(PO_4):Eu^{3+}$ [34], $NaBaPO_4$, $KBaPO_4$[35] prepared by solid state reaction and combustion method have been investigated. Eu^{3+} doped $NaCaPO_4$ by flux assisted method has been reported by Peng et. al [24]. However, Eu^{3+} doped $NaCaPO_4$ by molten salt and sol gel combustion synthesis technique is reported in the present manuscript for the first time to best of our knowledge.

In the present work, Eu^{3+} doped $NaCaPO_4$ phosphors were prepared by three different methods, such as solid-state reaction (SSR), molten salt synthesis (MSS) and sol-gel combustion (SGC). The relative structural, morphological and luminescence properties of the as-prepared phosphor were investigated. Eu^{3+} doped $NaCaPO_4$ phosphor was optimized by varying the dopant concentration through SGC method. Moreover, the effect of doping concentration on luminescent properties were discussed systematically.

2. Experimental work

Eu^{3+} doped $NaCaPO_4$ phosphors were synthesized by three different techniques such as solid-state reaction (SSR), molten salt (MS) and sol-gel combustion (SGC) methods. To synthesize all the samples, Na_2CO_3 (Fisher Scientific, 99.9%), $CaCO_3$ (Fisher Scientific, 99.9%), $(NH_3)_4HPO_4$ (Fisher Scientific, 99.9%) and Eu_2O_3 (Sigma-Aldrich, 99.99%) chemicals have been used as starting materials.

In our previous article, $NaCaPO_4$ phosphors have been synthesized via solid state reaction method [18]. The starting materials Na_2CO_3, $CaCO_3$, $NH_4H_2PO_4$ and Eu_2O_3 were used as precursors. All the materials were weighed in stoichiometric ratio and ground in an agate mortar for 1 h with ethanol as dispersing medium. Further, mixture was kept in an alumina crucible and placed inside the programmable muffle furnaces. The three steps heating procedure is used as, 185 °C for 2 h, 714 °C for 1 h and 950 °C for 3 h to synthesize required 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphors . Finally, the sample was naturally cooled to the room temperature and grounded again for further characterizations.

Accepted manuscript

ACCEPTED MANUSCRIPT

In molten salt synthesis, high purity eutectic mixture of NaCl and KCl in equal weight ratio were blended in the stoichiometric amount of $NaCaPO_4$ compound. The ratio of eutectic mixture and $NaCaPO_4$ was maintained as 5:1 weight percent. The stoichiometric ratio of the eutectic mixture and $NaCaPO_4$ were ground in agate mortar for an hour with ethanol as dispersed medium. The mixture was kept in an alumina crucible and placed in programmable muffle furnaces 800 °C for 3 h. Then the samples were naturally cooled to room temperature. Further, samples were washed several times with hot deionized water and dried in an oven at 150 °C for 24 h.

In the sol-gel combustion synthesis high purity $NaNO_3$ (sodium nitrate), $Ca(NO_3)_2{*}4H_2O$ (calcium nitrate tetrahydrate), $Eu(NO_3)_3$ (europium nitrate) and H_3PO_4 (phosphoric acid) were used as starting materials. Citric acid used as chelating agent and PEG-400 works as network former. A schematic flow chart for the preparation of the Eu^{3+} doped $NaCaPO_4$ by sol–gel combustion method is given in Fig. 1. Initially, the stoichiometric amount of $NaNO_3$, $Ca(NO_3)_2{*}4H_2O$, $(EuNO_3)_3$ and H_3PO_4 were dissolved in 40 ml deionized water. All these precursor solutions were mixed together and rigorously stirred with citric acid (metal to citric acid is in 1:5 molar ratio) for 30 min. Afterward, 2 ml PEG-400 was added in solution during stirring. The transparent and clear solution was obtained. Further, dropwise NH_4OH (ammonia solution, 25%) was added in solution to adjust the pH ~ 7. The final solution was evaporated slowly on the hot plate at 120 °C and formed a viscous thick brown gel. Subsequently, the temperature of hot plate was increased to 450 °C, leading to ignite of the gel, combustion process took place. The black fluffy porous material was obtained. Finally, the obtained material was ground in a mortar pestle and sintered at 800 °C for 3 h in a muffle furnace to obtain desired compound.

The phase purity of the synthesized samples was identified by x-ray diffraction (XRD) pattern using Bruker D8 Advanced powder X-ray diffractometer with a Cu-Kα radiation (λ =

Accepted manuscript

ACCEPTED MANUSCRIPT

1.5404 Å, 40 kV, 40 mA) source in the 2θ range of 20° to 70°. The surface morphology was observed via field emission scanning electron microscopy (FE-SEM; Tescan, MIRA II LMH). The excitation and emission spectra of the samples were recorded at room temperature by a Shimadzu RF-5301PC Spectrofluorophotometer fitted with a xenon lamp as the excitation source.

3. Result and Discussion

3.1. XRD and Rietveld refinement analysis

In order to obtain pure phase, $NaCaPO_4$ phosphors were synthesized via MSS and SGC methods initially at 600 °C and the corresponding XRD patterns are shown in Fig.2 (a) along with JCPDS data. The samples exhibit unknown diffraction peaks. Therefore, the diffraction patterns measured for the samples calcined at 800 °C via MSS and SGC method exhibit similar diffraction patterns synthesized by SSR and standard JCPDS pattern as shown in Fig. 2 (b and c). All the diffraction peaks of the samples are well matched with the standard JCPDS card no. 76-1456 of $NaCaPO_4$. Further, no unknown peaks were found, which confirm the formation of the single phase $NaCaPO_4$ with orthorhombic structure having space group $Pn2_1a$ (33). The strain (ε) and crystallite size (D) in the synthesized sample by different routes were estimated using Williamson- Hall (W-H) equation [36,37]:

$$\beta \cos\theta/\lambda = \frac{K}{D} + 4\varepsilon \sin\theta/\lambda$$

where, β is full-width half maximum (FWHM), θ = Bragg's diffraction angle, K= shape factor (K = 0.91), λ= x-ray wavelength (λ = 0.154 nm Cu-Kα radiation), D = average crystallite size and ε is the strain present in the sample. Assuming βcosθ/λ and 4sinθ as y and x-axis respectively, the straight line fitted curves were obtained as shown in Fig. 3 (a, b, and c). The slope of fitted line gives the strain (ε) and the inverse of the intercept (K/D) on y-axis gives crystallite size. The estimated strain (ε) value and crystallite size are represented in

Accepted manuscript

ACCEPTED MANUSCRIPT

Table 1. It has been observed that crystallite size reduces and simultaneously strain present in the sample increases while changing the synthesis routes from SSR to MSS and SGC [38,39].

The Rietveld analysis was performed for pure samples synthesized via SSR, MS and SGC routes. The peak shape was refined by pseudo-voigt function, and the background was described in the mode of linear interpolation between a set background points. The raw, fitted and difference results from the refinement of $NaCaPO_4$ are shown in Fig. 4 (a, b and c) for SSR, MSS and SGC synthesis, respectively. The observed and calculated patterns represent the good agreement with standard JCPDS data. The refinement parameters indicate that the diffraction pattern exhibits good agreement with the standard JCPDS data without anonymous peak. The calculated lattice parameters and goodness of fit (χ^2) are listed in Table 2. According to the refinement result for the crystal lattice of $NaCaPO_4$, there are three kinds of Na^+/Ca^{2+} sites and twelve sites of O^{2-} ions. This means that the activators Eu^{3+} ions could occupy the three Ca^{2+} cations site in the $NaCaPO_4$ sample. It is observed that the one P atom is surrounded by four O atoms and forms tetrahedral [PO_4] structure [16].

3.2 Morphological Analysis

The surface morphology of $NaCaPO_4$ synthesized via different routes are shown in Fig. 5. The FE-SEM image for the sample synthesized via SSR method, Fig. 5 (a), exhibits inhomogeneous, micron size particle with aggregate and irregular shape. Fig. 5 (b) represents the FE-SEM image of $NaCaPO_4$ samples as-prepared via SGC method. The sample exhibits irregular and melting morphology with huge agglomeration in micron size of particles as compared to the sample prepared via SSR and MSS methods, which may be due to sintering the sample at 800 ^oC [7]. Fig. 5 (c & d) represent the FE-SEM images for the sample synthesized via molten salt synthesis exhibit uniformly distributed spherical particles in the size of 4-5 micrometer range. In the MSS, eutectic mixture (NaCl- KCl) salt plays crucial

Accepted manuscript

ACCEPTED MANUSCRIPT

role in the formation of spherical shape particle. The mixture of eutectic composition is used to lower the liquid formation temperature at ~600 ºC and provides liquid phase ionic molten salt medium [40]. In melting state, K^+, Na^+ and Cl^- ions affect the crystallization and growth orientation of the particles and prevent from particle agglomerations [41].

3.3. Excitation and emission Spectra

Fig. 6. represents the excitation spectra measured by monitoring the emission peak at 595 nm for 1 mol% Eu^{3+} doped $NaCaPO_4$ synthesized by SGC, MSS and SSR method. The excitation spectra consist of a broad absorption band between 220 to 280 nm with several sharp peaks between 350 to 500 nm. A broadband with maxima at 256 nm attributed due to the charge transfer band (CTB). Moreover, CTB corresponds to the electronic transition from 2p orbital of O^{2-} to the empty 4f shell of the Eu^{3+} ion. It may be endorsed as a ligand to metal charge transfer (LMCT) transitions. Other sharp peaks are observed due to the 4f transitions of Eu^{3+} ions, assigned as $^7F_0 \rightarrow {}^5H_6$, $^7F_0 \rightarrow {}^5D_4$, $^7F_0 \rightarrow {}^5L_7$, $^7F_0 \rightarrow {}^5L_6$, $^7F_0 \rightarrow {}^5D_3$ and $^7F_0 \rightarrow {}^5D_2$ centered at wavelength 317, 360, 381, 392, 417 and 466 nm, respectively. The most intense sharp peak, $^7F_0 \rightarrow {}^5L_6$, at 392 nm confirms that the prepared phosphor may be effectively excited by n-UV LEDs in $NaCaPO_4$ host.

Fig. 7 represents the emission spectra of 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphors synthesized by SSR, MSS and SGC techniques under 392 nm excitation wavelength. All the emission spectra resemble the same band shape/profile with variation in intensities. The phosphor synthesized by SGC technique exhibit strongest emission relative to phosphor prepared by SSR and MSS methods [42–44]. Among all synthesis methods, the emission intensity of SGC synthesized phosphor is higher than others. This is due to the citric acid which helps in grain growth and improvement in crystallinity of as prepared phosphors. Solution based sol gel combustion method always helps to improve the crystallinity of

Accepted manuscript

ACCEPTED MANUSCRIPT

compounds which gives more intense emission [45,46]. This result is consistent with our XRD results in which SGC synthesized compound exhibit more intense diffraction peak. In MSS method, the spherical particle in micron range may enhance the PL emission intensity when compared to SSR method. Hence, the SGC method is more efficient method to synthesize $NaCaPO_4$: Eu^{3+} phosphors based on the emission intensity while comparing to other methods, which have been chosen for further investigation.

Fig. 8 shows the emission spectra of 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphor under 256, 381, 392 and 466 nm excitation wavelength. All the UV, n-UV and blue excited spectra exhibit similar emission lines at 580, 595, 622, 654 and 702 nm wavelength ascribed to $^5D_0 \rightarrow {}^7F_0$, $^5D_0 \rightarrow {}^7F_1$, $^5D_0 \rightarrow {}^7F_2$, $^5D_0 \rightarrow {}^7F_3$ and $^5D_0 \rightarrow {}^7F_4$ transitions of Eu^{3+} ions, respectively. Among all emission spectra, emission measured under n-UV excitation (392 nm) exhibit higher intensity confirm that the samples are efficient to excite under n-UV light. The intense emission peak centered at 595 nm ($^5D_0 \rightarrow {}^7F_1$) is attributed to a magnetic dipole (MD) transition and an emission peak at 622 nm ($^5D_0 \rightarrow {}^7F_2$) ascribed to electric dipole (ED) transition. The emission intensity of $^5D_0 \rightarrow {}^7F_1$ at 595 nm is 1.33 times higher than that of $^5D_0 \rightarrow {}^7F_2$ at 622 nm suggesting a higher occupancy of Eu^{3+} ions in the symmetric environment. Therefore, the activator Eu^{3+} (1.07Å) occupies the Ca^{2+} (1.12Å) site with an inversion symmetry. The transition $^5D_0 \rightarrow {}^7F_3$ (654 nm) was strictly forbidden and exhibit weak intensity due to J-mixing effect between multiplets, and $^5D_0 \rightarrow {}^7F_4$ (702 nm) correspond to induced electric dipole transition.

Fig. 9 represent the SGC synthesized emission spectra of $NaCaPO_4$: xEu^{3+} (x = 0.5, 1, 1.5, 2 and 3 mol%) phosphors, under 392 nm excitation wavelength. It is observed that increase in Eu^{3+} doping concentration leads to the enhancement in the emission intensity up to 1 mol%. Beyond this concentration of Eu^{3+} ion, the emission intensity starts decreasing, which may due to concentration quenching phenomena [9,47]. The concentration quenching

Accepted manuscript

ACCEPTED MANUSCRIPT

phenomenon is a result induced by the non-radiative energy transfer between the two adjacent Eu^{3+} ions. Thus, Blasse suggested that the activator-activator pair distance (critical distance) can be calculated by the following formula [48]:

$$R_c = 2\left(\frac{3V}{4\pi x_c N}\right)^{1/3}$$

where V is the unit cell volume, X_c = activator concentration (mol) and N is the number of the site in the host unit cell. In the case of 0.01 mol% Eu^{3+} doped $NaCaPO_4$ phosphor, the parameters, V = 1011.73 $Å^3$ and N = 4, are used to calculate the critical distance and found to be R_c ~ 36.41 Å [18]. Non-radiative energy transfer occurs by the exchange interaction, a radiation re-absorption or a multipolar interaction. The exchange interaction is generally responsible for the energy transfer for forbidden transitions and the critical distance should be about 5 Å. Since, the distance between adjacent Eu^{3+} ions is larger than 5 Å, so it may confirm that the interaction is multipolar [36]. At high activator concentration, critical distance decreases that leads to cross relaxation between Eu^{3+} ions which in turn decrease in PL intensity. From the inset plot of the Fig. 9 it is confirmed that the 1 mol% Eu^{3+} is the optimized concentration for $NaCaPO_4$ when synthesized via SGC routes.

3.4. CIE chromaticity coordinates

In order to reveal the PL emission color of the phosphor, calculation of CIE chromaticity coordinates are essentially required. CIE color coordinates were calculated based on the emission spectra under 392 nm excitation wavelength for 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphors synthesized by different methods. The calculated color coordinates are found to be (0.621, 0.377), (0.620, 0.378) and (0.622, 0.376) for different phosphors synthesized via SGC, MSS and SSR methods. The color coordinates (x, y) were located in the orange-red region for all the phosphor synthesized via different synthesis routes. From the

Accepted manuscript

ACCEPTED MANUSCRIPT

Fig. 10 it is confirmed that the color coordinates does not change with the synthesis routes as there is no variation in PL emission band with the synthesis routes.

4. Conclusions

In summary, Eu^{3+} activated $NaCaPO_4$ phosphor have been prepared by different routes such as SSR, MSS and SGC. The comparative study of the structural, morphological and luminescence properties of as-prepared phosphor have been established. The phase purity of $NaCaPO_4$ has been confirmed by the XRD and Rietveld refinement. All the phosphor samples exhibit orange-red emission under the near UV and blue excitations. The intense emission was observed for 1 mol% Eu^{3+} doped $NaCaPO_4$ sample synthesized by SGC method and this sample has been considered as optimized. The CIE chromaticity coordinates were calculated and concluded that all the samples are located in orange-red region. The above results indicate that SGC is the best route to synthesize the efficient Eu^{3+} doped $NaCaPO_4$ phosphor and also could be a potential candidate for display and solid state lighting applications.

Accepted manuscript

Acknowledgements:

One of the authors (Dr. M. Jayasimhadri) would like to thank DST-SERB, Govt. of India for sanctioning the research project (No. SB/FTP/PS-082/2014, dt. 02/03/2015). This work was also partially supported by the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (NRF–2015–056886).

References:

[1] J.H. Oh, S.J. Yang, Y.R. Do, Light Sci. Appl. (2014) 3.

[2] X. Li, J.D. Budai, F. Liu, J.Y. Howe, J. Zhang, X. Wang, Z. Gu, C. Sun, R.S. Meltzer,

ACCEPTED MANUSCRIPT

Z. Pan, Light Sci. Appl. (2013) 1–8.

[3] K. Li, J. Fan, X. Mi, Y. Zhang, H. Lian, M. Shang, J. Lin, Inorg Chem 53 (2014) 12141–12150.

[4] J. Li, H. Yan, F. Yan, Optik (Stuttg). 127 (2016) 4541–4544.

[5] L. Mishra, A. Sharma, A.K. Vishwakarma, K. Jha, M. Jayasimhadri, B. V Ratnam, K. Jang, A.S. Rao, R.K. Sinha, J. Lumin. 169 (2016) 121–127.

[6] A.K. Vishwakarma, K. Jha, M. Jayasimhadri, A.S. Rao, K. Jang, B. Sivaiah, D. Haranath, J. Alloys Compd. 622 (2015) 97–101.

[7] B.K.G. and K.K.C. Subhajit Saha, Swati Das, Uttam Kumar Ghorai, Nilesh Mazumder, Dalt. Trans. 42 (2013) 12965–12974.

[8] B. Han, P. Li, J. Zhang, H. Shi, J. Lumin. 155 (2014) 15–20.

[9] A.K. Vishwakarma, M. Jayasimhadri, J. Lumin. 176 (2016) 112–117.

[10] L.T. Francis, P.P. Rao, M. Thomas, S.K. Mahesh, V.R. Reshmi, V.D.S. Thampi, Mater. Lett. 81 (2012) 142–144.

[11] D. Haranath, S. Mishra, S. Yadav, R.K. Sharma, L.M. Kandpal, N. Vijayan, M.K. Dalai, G. Sehgal, V. Shanker, Appl. Phys. Lett. 101 (2012) 1–6.

[12] Z. Wei, J. Guan, J. Song, Y. Xu, Superlattices Microstruct. 88 (2015) 308–314.

[13] J. Gou, J. Wang, B. Yu, D. Duan, S. Ye, S. Liu, Mater. Res. Bull. 64 (2015) 364–369.

[14] R.Y. Yang, C.T. Pan, K.H. Chen, C.Y. Hung, Opt. Mater. (Amst). 35 (2013) 2183–2187.

[15] Z. Liang, F. Mo, X. Zhang, L. Zhou, C. Xu, Ceram. Int. 40 (2014) 7501–7506.

[16] N. Guo, Y. Song, H. You, G. Jia, M. Yang, K. Liu, Eur. J. Inorg. Chem. (2010) 4636–

Accepted manuscript

ACCEPTED MANUSCRIPT

4642.

[17] L. Shi, Y. Huang, H.J. Seo, J. Phys. Chem. A 114 (2010) 6927–6934.

[18] and H.S.L. B.V. Ratnam, M. Jayasimhadriz, Kiwan Jangw, J. Am. Ceram. Soc. 93 (2010) 3857–3861.

[19] G. Li, X. Xu, C. Peng, M. Shang, D. Geng, J. Chen, J. Lin, Opt. Express 19 (2011) 3020–3028.

[20] B. Kumar Grandhe, V. Rao Bandi, K. Jang, H.-S. Lee, D.-S. Shin, S.-S. Yi, J.-H. Jeong, Ceram. Int. 38 (2012) 6273–6279.

[21] C.C. Wu, K.B. Chen, C. Sen Lee, T.M. Chen, B.M. Cheng, Chem. Mater. 19 (2007) 3278–3285.

[22] S. Som, S. Das, S. Dutta, H.G. Visser, K. Pandey, RSC Adv. (2015) 70887–70898.

[23] P. Godlewska, A. Matraszek, L. Macalik, K. Hermanowicz, M. Ptak, P.E. Tomaszewski, J. Hanuza, I. Szczygieł, J. Alloys Compd. 628 (2015) 199–207.

[24] Y. Peng, Y. Su, R. Yang, Mater. Res. Bull. 48 (2013) 1946–1951.

[25] B. Yan, F. Lei, J. Alloys Compd. 507 (2010) 460–464.

[26] M.S. Synthesis, J. Am. Ceram. Soc. 51 (2012) 49–51.

[27] T. Igarashi, M. Ihara, T. Kusunoki, K. Ohno, T. Isobe, M. Senna, Appl. Phys. Lett. 76 (2000) 1549.

[28] H. Chander, Mater. Sci. Eng. R Reports 49 (2005) 113–155.

[29] Y.Y. Meng, M.H. He, Q. Zeng, D.L. Jiao, S. Shukla, R. V Ramanujan, Z.W. Liu, J. Alloys Compd. 583 (2014) 220–225.

[30] L. Shi, C. Zeng, Y. Jin, N. Tsubaki, Catal. Sci. Technol. (2012) 2569–2577.

Accepted manuscript

ACCEPTED MANUSCRIPT

[31] H. Pan, X. Li, J. Zhang, L. Guan, H. Su, F. Teng, 155 (2015) 106–108.

[32] Y.S. Tang, S.F. Hu, C.C. Lin, N.C. Bagkar, R.S. Liu, Appl. Phys. Lett. 90 (2007) 2–5.

[33] Y. Zhang, Z. Mao, D. Wang, J. Zhao, Mater. Res. Bull. 67 (2015) 1–4.

[34] Q. Long, Y. Gao, Y. Huang, S. Liao, B. Song, W. Wu, J. Cai, Mater. Lett. 160 (2015) 436–439.

[35] T. Wanjun, C. Donghua, J. Am. Ceram. Soc. 92 (2009) 1059.

[36] A.K. Vishwakarma, K. Jha, M. Jayasimhadri, B. Sivaiah, B. Gahtori, D. Haranath, Dalt. Trans. 44 (2015) 17166–17174.

[37] M. Chandrasekhar, H. Nagabhushana, Y.S. Vidya, K.S. Anantharaju, S.C. Sharma, H.P. Nagaswarupa, J. Mol. Catal. A 409 (2015) 26–41.

[38] H. Wang, Y. Liu, M. Li, H. Huang, H.M. Xu, R.J. Hong, H. Shen, Phys. Chem. Chem. Phys. 4 (2010) 1166–1169.

[39] M.K. Mahata, S.P. Tiwari, S. Mukherjee, K. Kumar, V.K. Rai, J.Opt. Soc. Am.B 31 (2014) 1814–1821.

[40] G. Hu, W. Li, G. He, J. Wang, Y. Zhao, Y. Zhang, Ceram. Int. 42 (2016) 7852–7856.

[41] X. Wang, J. Li, P. Shi, W. Guan, H. Zhang, Opt. Mater. (Amst). 46 (2015) 432–437.

[42] Y.H. Liu, J.H. Hao, W.D. Zhuang, Y.S. Hu, J. Phys. D-Applied Phys. 42 (2009).

[43] M. Kim, M. Kobayashi, H. Kato, M. Kakihana, Opt. Photonics J. 3 (2013) 13–18.

[44] Y. Chen, J. Wang, M. Gong, Q. Su, J. Solid State Chem. 180 (2007) 1165–1170.

[45] H. Sameie, R. Salimi, A.A. Sarabi, A. Asghar, S. Alvani, S. Nargesian, H.E. Mohammadloo, Y. Ebrahimi, 47 (2013) 685–687.

Accepted manuscript

ACCEPTED MANUSCRIPT

[46] A. Poddar, S.C. Gedam, S.J. Dhoble, J. Lumin. 158 (2015) 188–196.

[47] B. V Ratnam, M. Jayasimhadri, K. Jang, Spectorchimica Acta Part A Mol. Biomol. Spectrosc. 132 (2014) 563–567.

[48] G. Blasse, Philips Res. 24 (1969) 131–144.

Accepted manuscript

ACCEPTED MANUSCRIPT

Fig.1: Synthesis procedure for $NaCaPO_4$ phosphor via SGC method.

Fig. 2(a): X-ray diffraction pattern of $NaCaPO_4$ synthesized via SSR, MSS and SGC methods calcined at 950, 600 and 600 °C respectively. **Fig. 2(b) :** X-ray diffraction pattern (Intensity vs 2θ) of $NaCaPO_4$ synthesized via SSR, MSS and SGC methods calcined at 950, 800 and 800 °C, respectively. **Fig. 2(c) :** X-ray diffraction pattern (Log I vs 2θ) of $NaCaPO_4$ synthesized via SSR, MSS and SGC methods calcined at 950, 800 and 800 °C respectively.

Fig.3: W-H plot of $NaCaPO_4$ synthesized via (a) SSR, (b) MSS and (c) SGC methods.

Fig.4: Rietveld refinement of $NaCaPO_4$ synthesized via (a) SSR, (b) MSS and (c) SGC methods.

Fig.5: FE-SEM images of $NaCaPO_4$ phosphors synthesized by (a) SSR, (b) SGC and (c & d) MSS.

Fig.6: Excitation spectra of 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphors synthesized via different routes.

Fig. 7: Emission spectra of 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphors synthesized via different routes. (Inset: Gaussian fit for $^5D_0 \rightarrow {}^7F_0$ transition)

Fig.8: Emission spectra of Eu^{3+} doped $NaCaPO_4$ phosphors synthesized via SGC method under different excitations.

Fig.9: Emission spectra of $NaCaPO_4$ phosphors for different Eu^{3+} concentrations synthesized via SGC method.

Fig.10: CIE color coordinates of 1 mol% Eu^{3+} doped $NaCaPO_4$ phosphors synthesized via different methods.

Accepted manuscript

ACCEPTED MANUSCRIPT

Table 1: Crystallite size and strain present in samples ($NaCaPO_4$) prepared via different routes.

Synthesis methods	Crystallite size (nm) W-H plot	Strain ($X10^{-4}$)
SSR	60.97	1.229
MSS	46.72	2.072
SGC	35.58	2.825

Table 2: The unit cell refined parameters along with the standard JCPDS data.

Parameters	**Standard data JCPDS (76-1456)**	**$NaCaPO_4$ SSR**	**$NaCaPO_4$ MSS**	**$NaCaPO_4$ SGC**
a (Å)	20.3970	20.3888	20.4032	20.4819
b (Å)	5.4120	5.4059	5.4038	5.3953
c (Å)	9.1610	9.1601	9.1696	9.1554
$\alpha = \beta = \gamma$	90°	90°	90°	90°
GOF (χ^2)	-	1.28	1.22	1.38
Cell Volume ($Å^3$)	1011.27	1009.62	1010.97	1011.73

Accepted manuscript

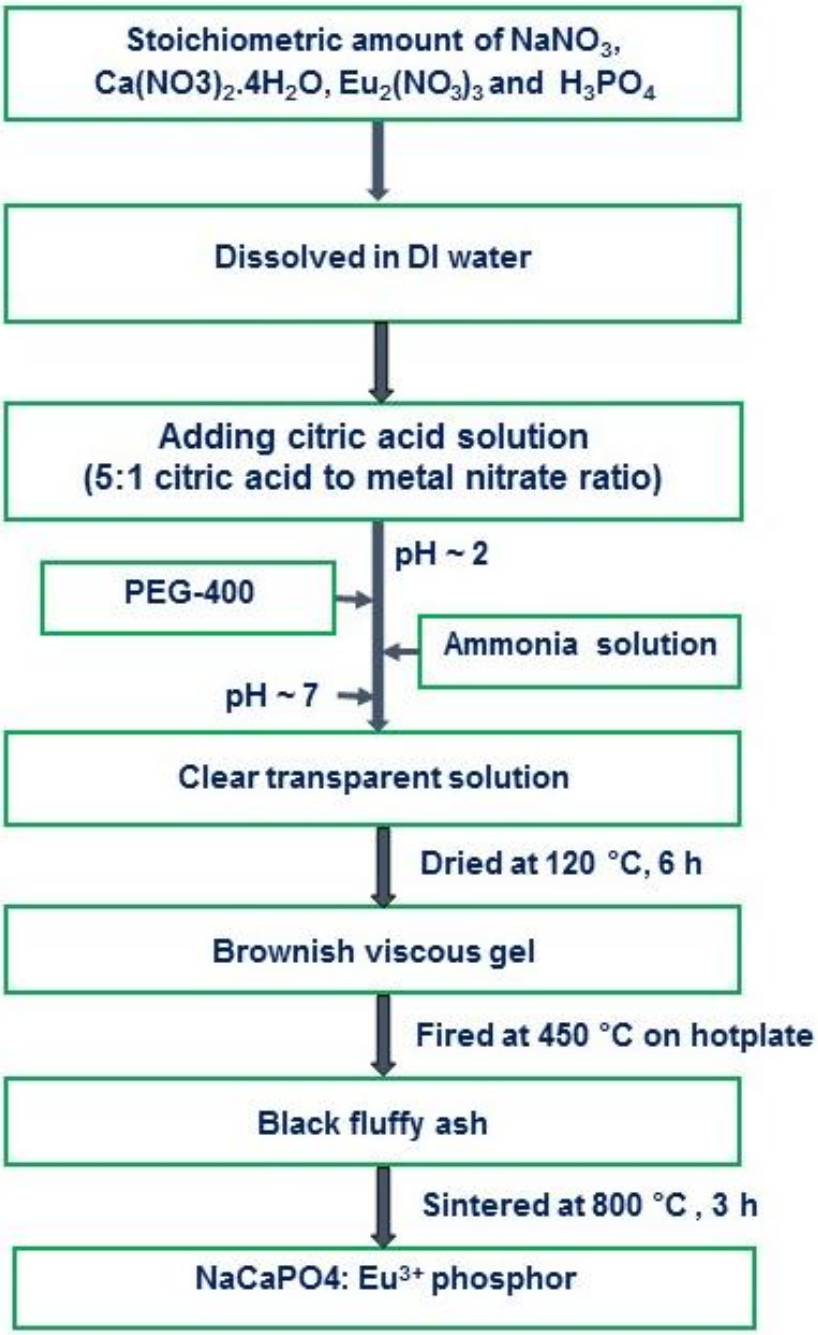

Stoichiometric amount of $NaNO_3$, $Ca(NO3)_2.4H_2O$, $Eu_2(NO_3)_3$ and H_3PO_4
Dissolved in DI water
Adding citric acid solution (5:1 citric acid to metal nitrate ratio)
pH ~ 2
PEG-400
Ammonia solution
pH ~ 7
Clear transparent solution
Dried at 120 °C, 6 h
Brownish viscous gel
Fired at 450 °C on hotplate
Black fluffy ash
Sintered at 800 °C , 3 h
NaCaPO4: Eu^{3+} phosphor

Fig. 1

ACCEPTED MANUSCRIPT

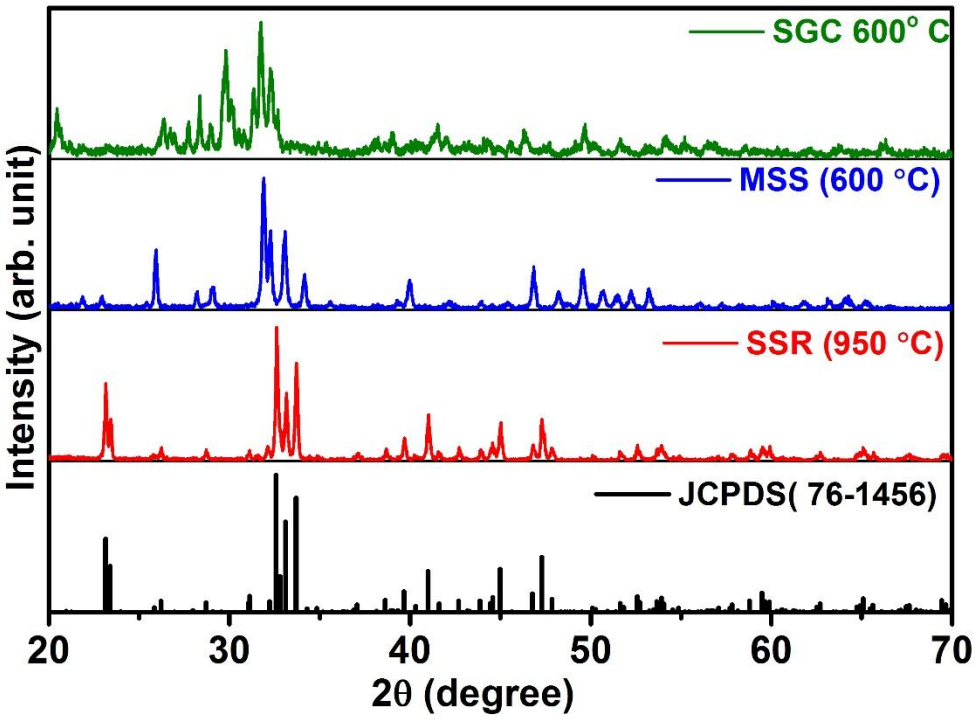


Fig. 2(a)

Accepted manu

ACCEPTED MANUSCRIPT

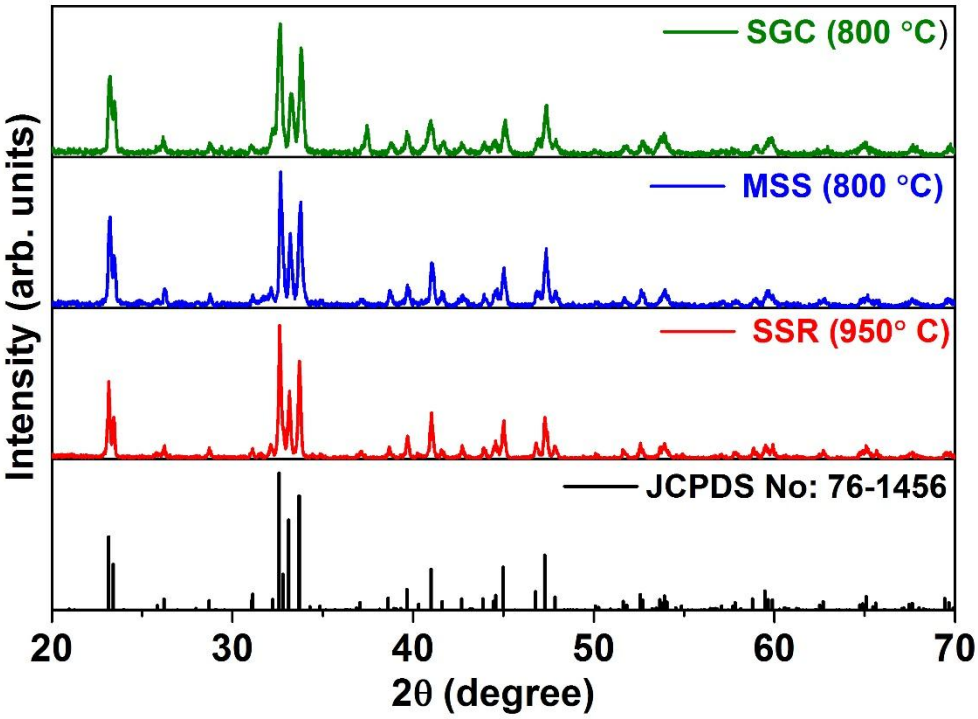


Fig. 2(b)

Accepted manu

ACCEPTED MANUSCRIPT

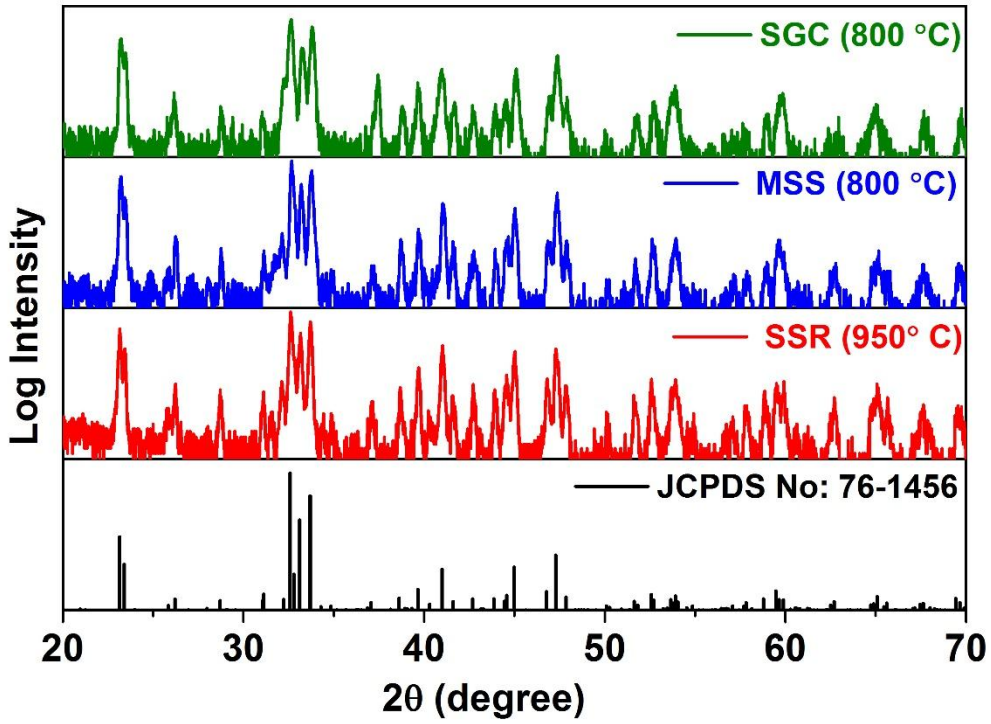


Fig. 2(c)

Accepted manu

ACCEPTED MANUSCRIPT

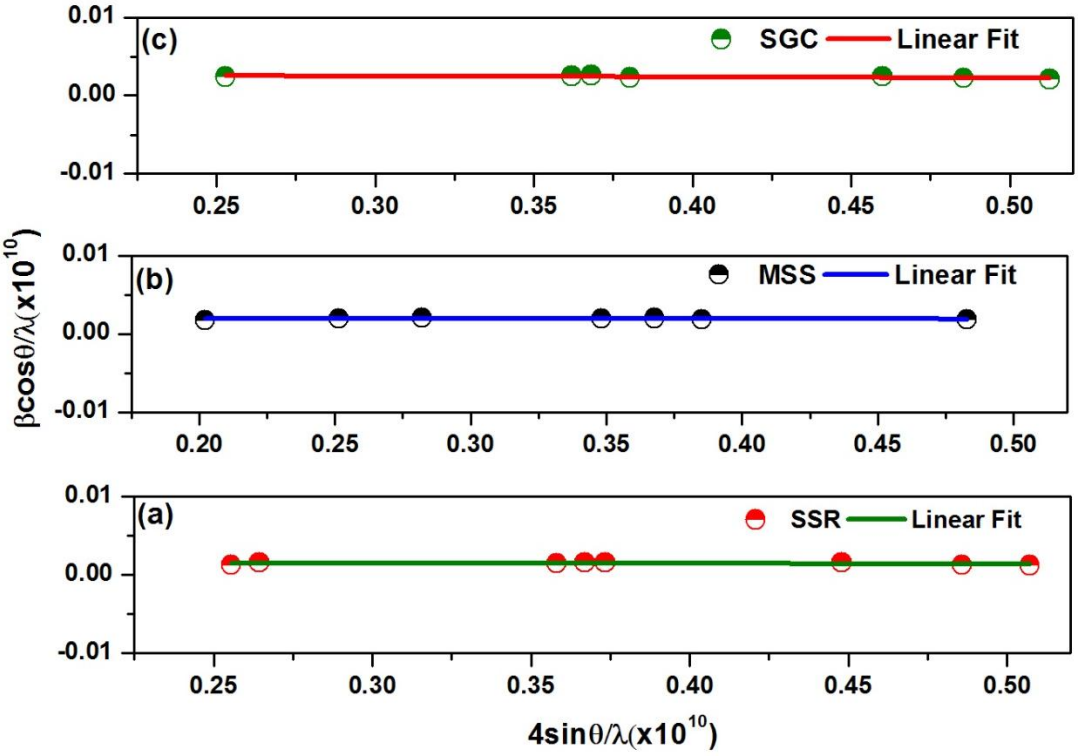


Fig. 3

Accepte

ACCEPTED MANUSCRIPT

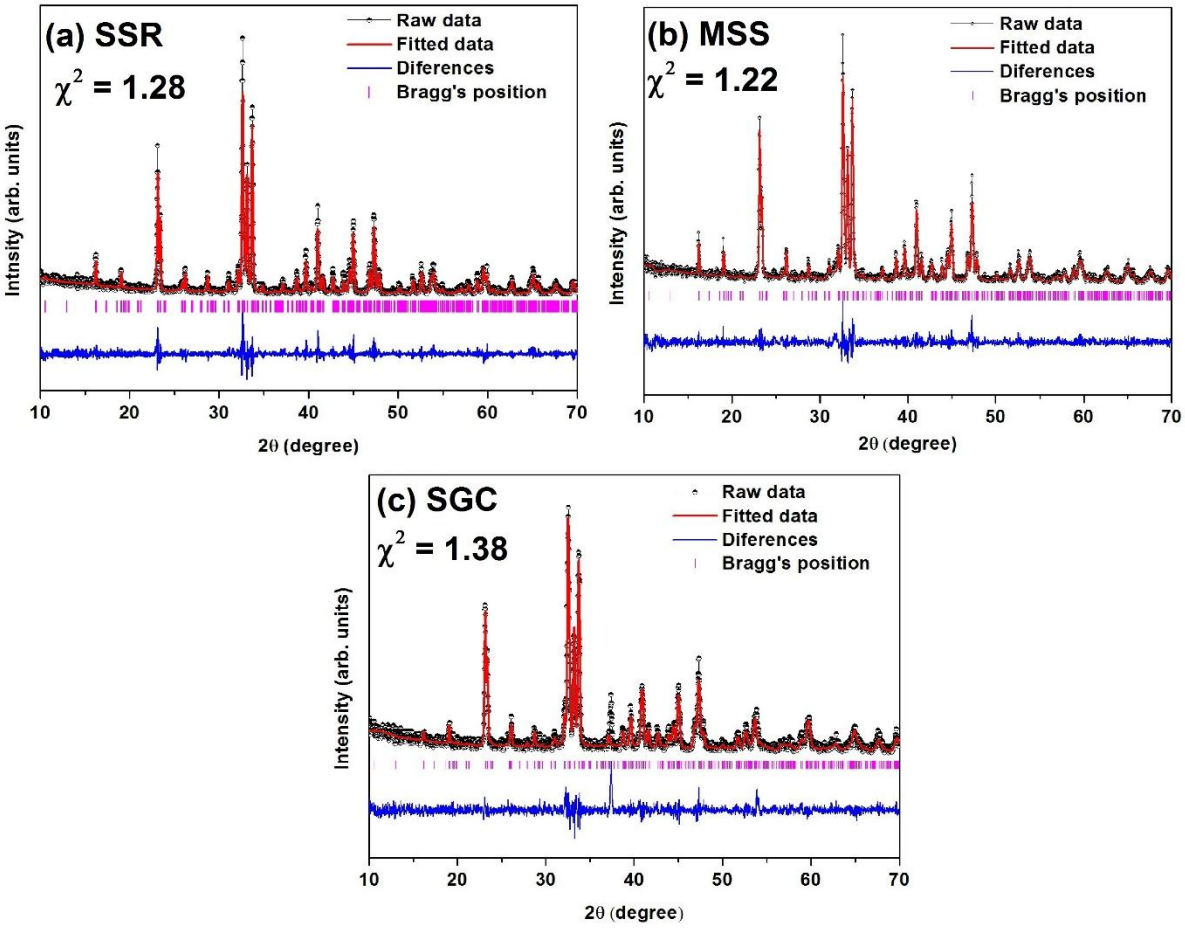


Fig. 4

Accepted ma

ACCEPTED MANUSCRIPT

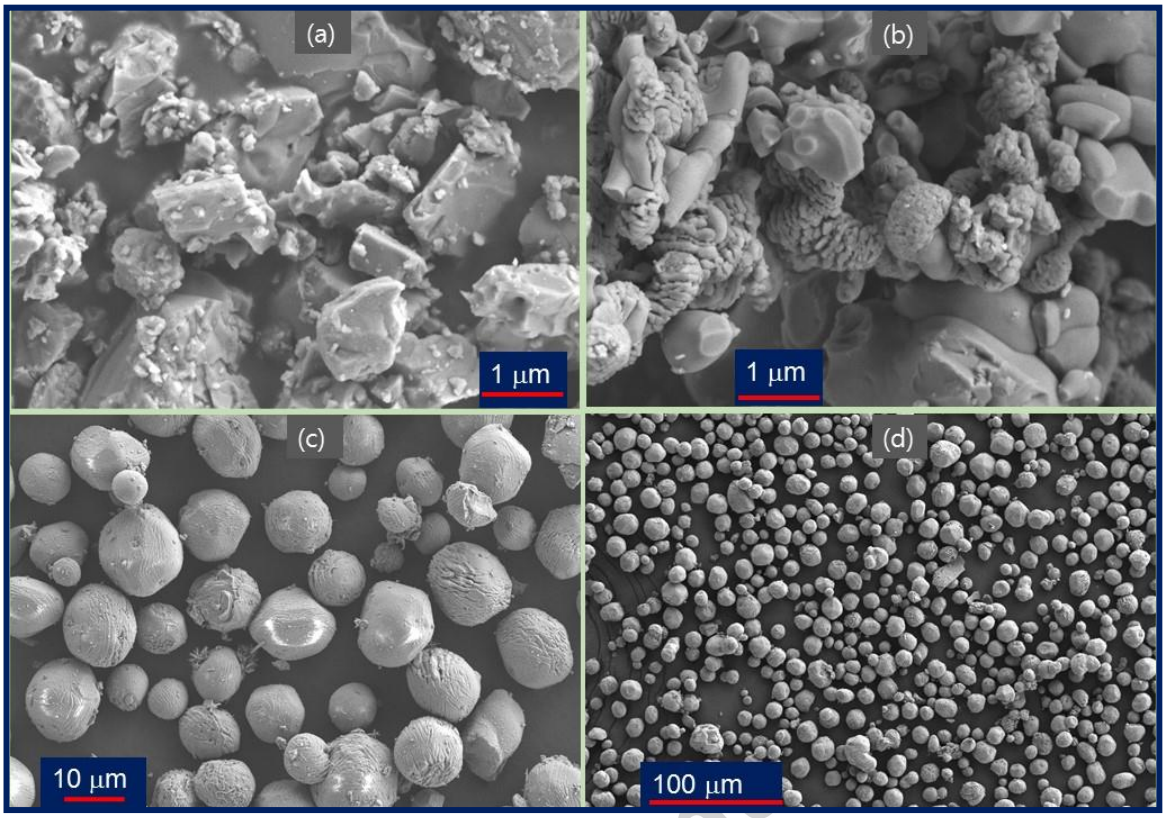


Fig. 5

Accepted manuscript

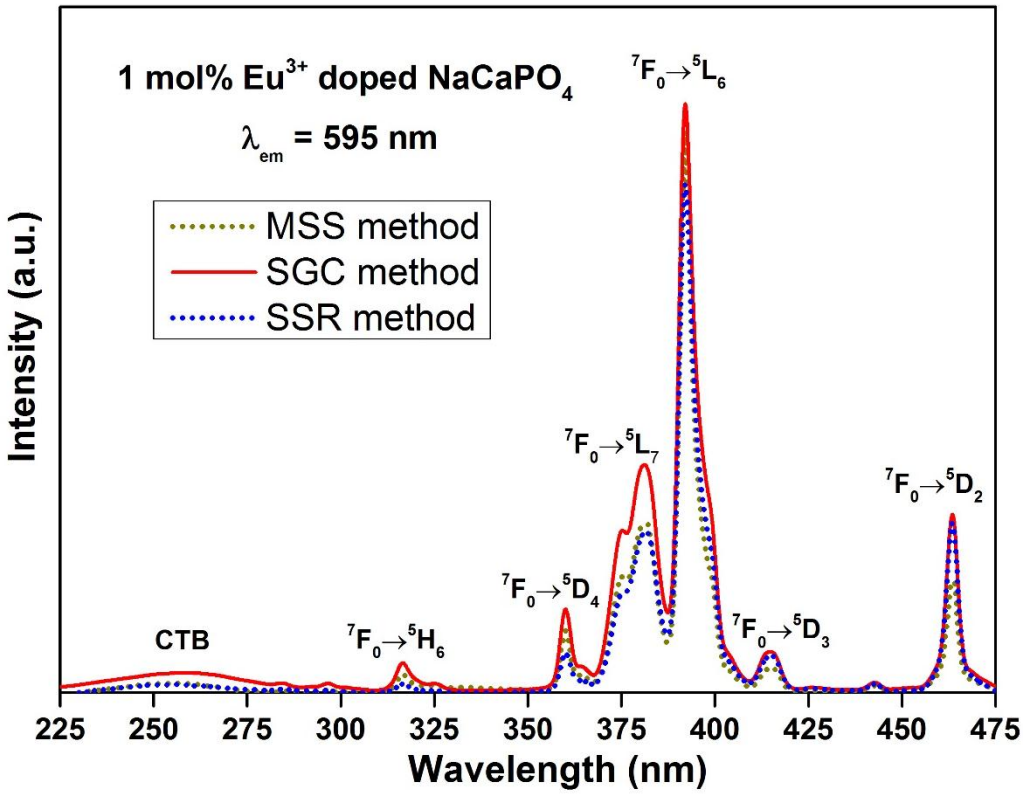


Fig. 6

Accepted manuscript

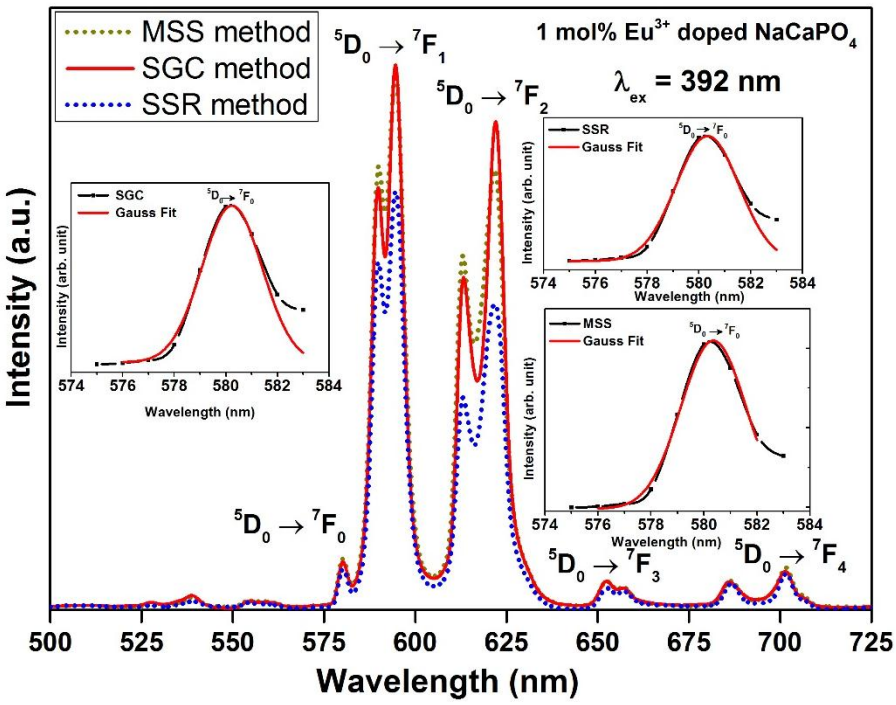


Fig. 7

Accepted manuscript

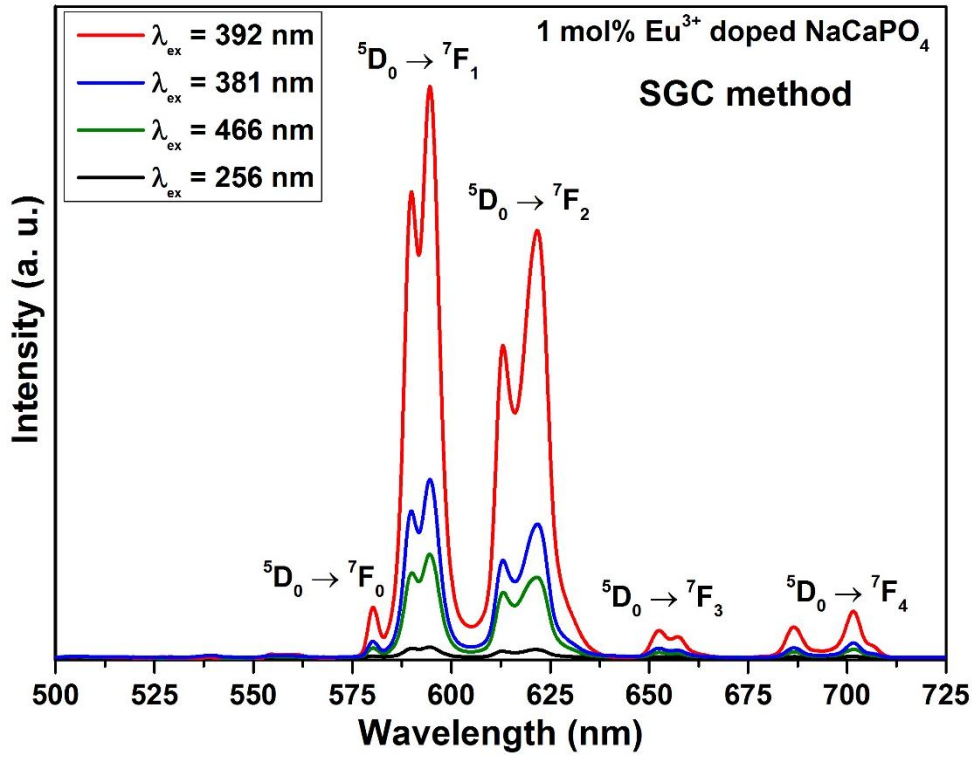


Fig. 8

Accepted manuscript

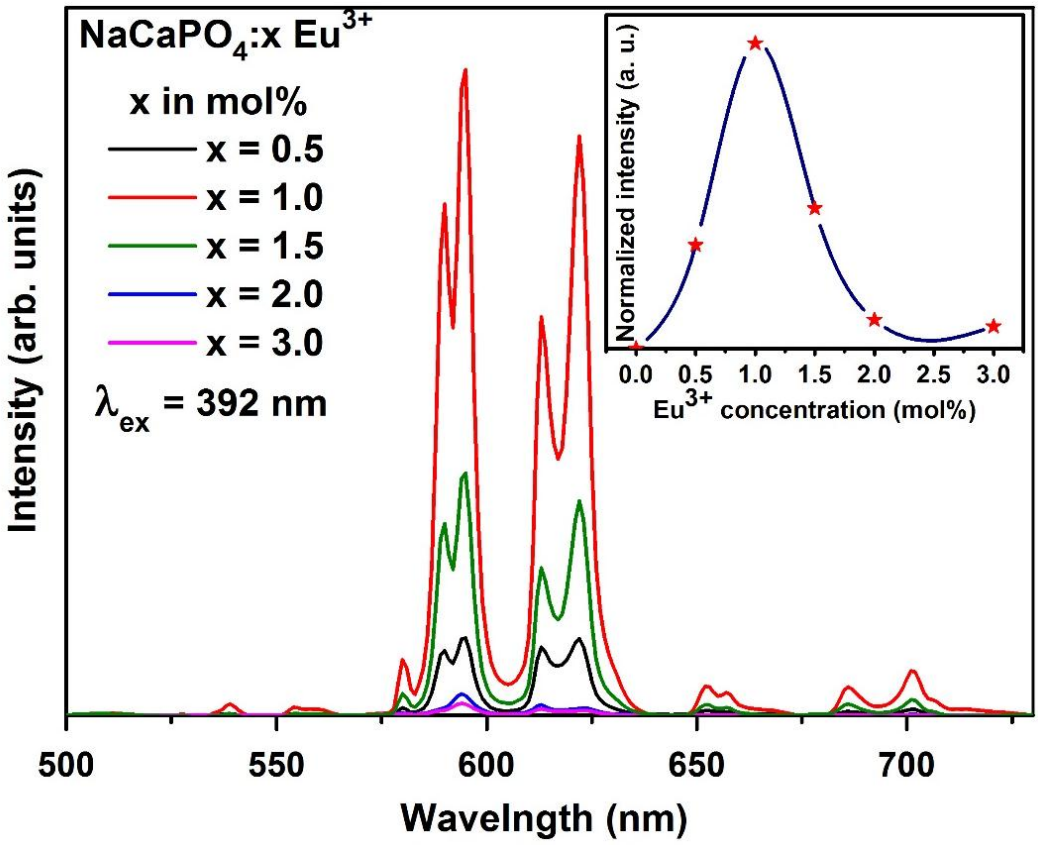

NaCaPO4:x Eu3+
x in mol%
x = 0.5
x = 1.0
x = 1.5
x = 2.0
x = 3.0
λex = 392 nm
Intensity (arb. units)
Wavelngth (nm)
500
550
600
650
700
Normalized intensity (a. u.)
Eu3+ concentration (mol%)
0.0
0.5
1.0
1.5
2.0
2.5
3.0

Fig. 9

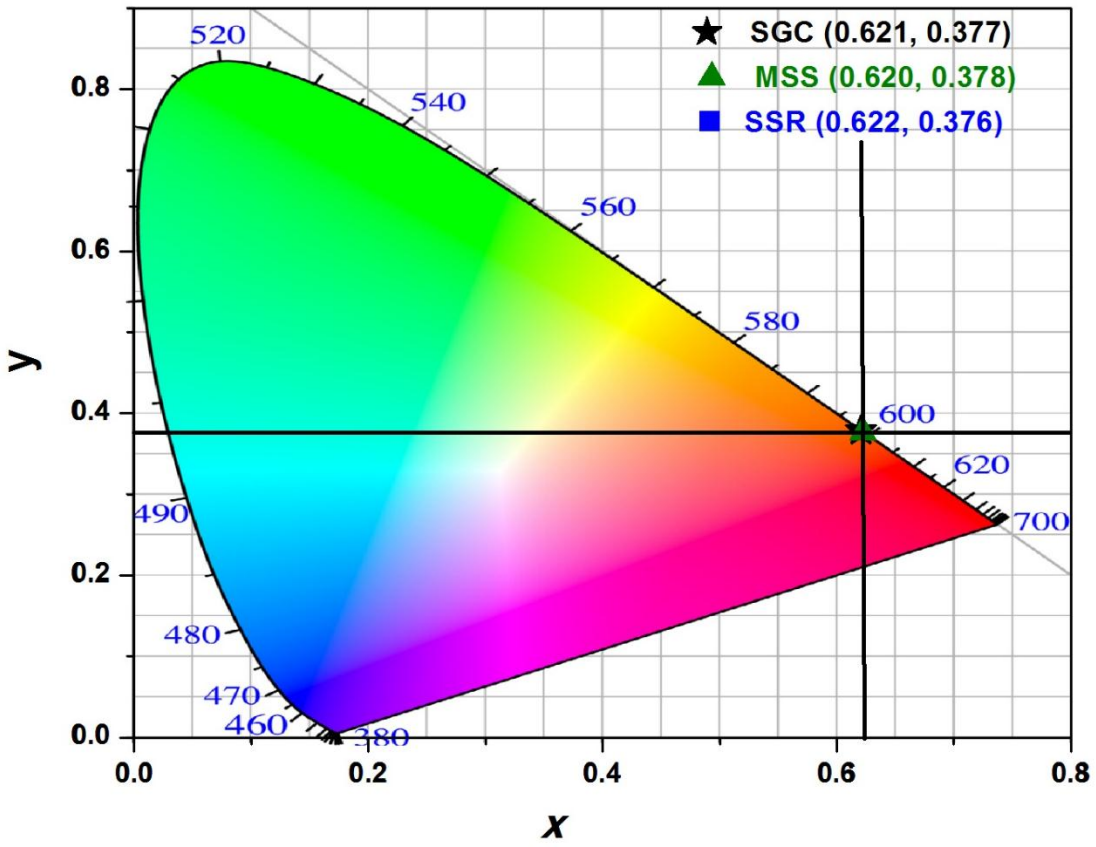

SGC (0.621, 0.377)
MSS (0.620, 0.378)
SSR (0.622, 0.376)
520
540
560
580
600
620
700
490
480
470
460
380
y
x
0.0
0.2
0.4
0.6
0.8

Fig. 10

ISSN 2466-2232
Online ISSN 2466-2100

Pulse TIG welding: Process, Automation and Control

P. K. Baghel*,† and D. S. Nagesh*

*Department of Mechanical Engineering, Delhi Technological University, Bawana Road, New Delhi 110042, India

†Corresponding author : pushpbaghel.dtu@gmail.com

(Received May 31, 2016 ; Revised September 26, 2016 ; Accepted October 19, 2016)

Abstract

Pulse TIG (Tungsten Inert Gas) welding is often considered the most difficult of all the welding processes commonly used in industry. Because the welder must maintain a short arc length, great care and skill are required to prevent contact between the electrode and the workpiece. Pulse TIG welding is most commonly used to weld thin sections of stainless steel, non-ferrous metals such as aluminum, magnesium and copper alloys. It is significantly slower than most other welding techniques and comparatively more complex and difficult to master as it requires greater welder dexterity than MIG or stick welding. The problems associated with manual TIG welding includes undercutting, tungsten inclusions, porosity, Heat affected zone cracks and also the adverse effect on health of welding gun operator due to amount of tungsten fumes produced during the welding process. This brings the necessity of automation. Hence, In this paper an attempt has been made to build a customerized setup of Pulse TIG welding based on through review of Pulse TIG welding parameters. The cost associated for making automated TIG is found to be low as compared to SPM (Special Purpose machines) available in the market.

Key Words : Pulse TIG welding, Automation, Review, TIG variants

Introduction

The Tungsten Inert Gas (TIG) welding is a welding process that uses a non-consumable tungsten electrode and an inert gas shield to protect the electrode, arc column and weld pool[1] as illustrated in Fig. 1. The weld arc acts as a heat source only and the welding engineer has a choice of whether or not to add a filler wire. Argon or helium is used as shielding gas to protect the weld from atmospheric oxidation. (V-I) characteristics of drooping type is used as power source. In the domain of arc welding processes TIG welding continues its apex position due to its versatility and flexibility in adaptation. The superior weld quality obtained in TIG weldments differentiates the TIG process in comparison with other competing and emerging joining process.

The modified version of TIG welding process is pulse TIG welding process in which pulse current alternates between a low level and peak level. Heating and fusion take place during peak current period and weld pool cooling. Solidification takes place during pulses as the heat is dissipated in the job during the low level current period. This current pulsing leads to intermittent melting along the joint seam giving a series of discrete melt spots, which overlap each other[2] . Current pulse regulator also adjust for the number of pulse per second and the duration of time spent at the peak current level which leads to control heat input and thus improves

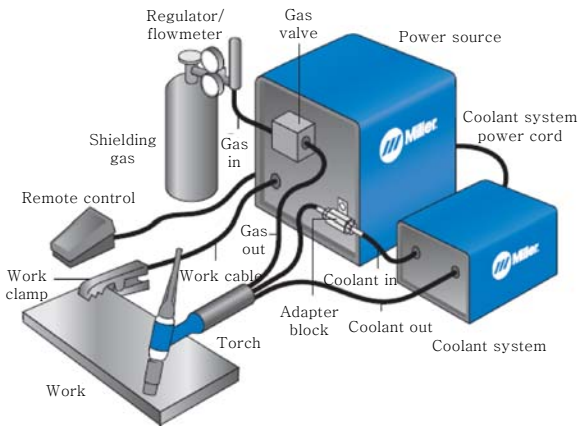


Fig. 1 TIG set-up[1]



Fig. 2 Pulse TIG welding machine

weld profile. DCEN is preferred current supply. The machine is shown in Fig. 2.

Pulse current characteristics are readily used to control the weld pool formation. High heat accelerates the weld pool formation which reduces the properties as gases from surrounding atmosphere like hydrogen readily diffused in weld pool inducing stress leading to cracks.

Due to controlled heat input, distortion and warpage are minimized specially for dissimilar thickness metal joining. It is also used to join dissimilar metals of same thicknesses such as copper to stainless steel, stainless steel to mild steel.

1. Machine parameters

Finding Optimum condition of Pulse TIG welding parameters are useful in metallurgical advantages of weldments. Literature frequently revealed improvement in grain refinement of fusion zone grain size and structure, reduced width of heat affected zone, control segregation because of current pulsing parameters[3]. The control heat input obtained due to current pulsing decreases the distortion and warpage which helps in improvement of mechanical properties[4,5]. From the recent literature available it is evident that the welding speed and wire feed rate are not considered as main parameters. The welding speed is proved significant parameter in controlling weld bead characteristics. Faster the welding speed, faster the cooling rate, smaller the grains formed resulting in maximum tensile properties[6,7]. The automation in welding process is required for superior weld quality at faster speeds. The automation in joining process is the main demand in the industry, but they have certain limitations as compared to semi-automatic system like

- High initial cost: Automated welding systems require higher initial investment. The amount required for setting up automation is estimated around $175,000 to $250,000 whereas semiautomatic systems often start around $30,000[8].
- Zero flexibility: The flexibility of a machine has an inverse relationship with the degree of automation. While semiautomatic machines can be moved from one part to the next, the automatic systems can only make a dedicated and pre-programmed manufacturing operation.
- Longer lead time: Setting and implementing automated welding system requires a longer lead time for reaching full-scale production. If a company needs to begin welding parts immediately, semiautomatic machines takes 4 to 8 weeks to deliver whereas fully automatic system takes at least 20 weeks and therefor takes longer time to outweigh the initial lead cost[9].
- Requirement of expertise in employees: Employees requires training programme for interacting with new robotic equipment. This normally takes time and financial output.
- High preventive maintenance: High preventive maintenance is required in automatic system because of complexity as compared to semiautomatic machines. The maintenance program includes cleaning and lubricating the machine, calibrating the controls and power supply, and replacing consumables.
- Consideration of product life cycle: Automation system does not suite for elements which follow unpredictable pattern of introduction, growth, maturity and decline in product life cycle. Whereas semiautomatic machines are readily used for such type of products. Batteries, sensors, fuel filters, medical components, nuclear devices, pipe fittings, transformer cores etc are products manufactured by semiautomatic systems.

2. Automation in pulse TIG welding machines

These are semi-automatic machines built for special purpose. They have several benefits which includes consistency in quality welds, greater repeatability and lowered production costs, fewer scrapped parts, increases return on investment and faster speed per cycle. karpagaraj et al.[10] used automatic setup for welding titanium alloy in which torch movement is automated using single axis mechanism. Tadayuki suggested to have built the automatic setup built having torch position specifically kept at 40° to yield good results[11]. The block diagram of automation in TIG welding has been given in Fig. 3.

An automatic setup is built to weld dissimilar metals having thickness greater than 6.35mm. To improve the quality of weld joint automation is essential especially

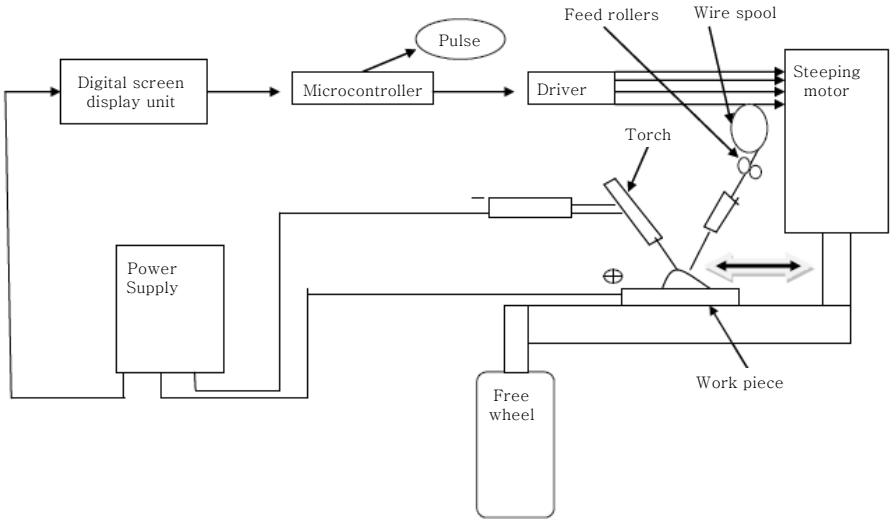


Fig. 3 Automatic TIG setup

where thickness of welding sections are greater than t > 5[12]. Also adverse effect on operator health due to amount of tungsten fumes, safety of operator and costly availability of SPM in the market are some of the reasons which urged to make automation in TIG welding process.

3.1 Automatic Torch movement

Automating the torch motions decreases the error potential which means decreased scrap and rework. The setup consists of trolley, pencil torch, track, fixture for holding pencil torch. The Weldcraft WP-9P, Light-Duty Front Loading air cooled TIG Pencil Torch for fixture mounted TIG applications is used which has following specifications.

WP-9P Specifications:

- 125 Amp DC, 100 Amp AC (60% duty cycle)
- Electrode capacity .020" (0.5mm) -- 1/8" (3.2mm)
- Torch body weight 3.oz (85.0g)
- Torch length 6" (152mm)
- Handle diameter 3/4" (19.0mm)

Roshan et al.[13] uses single axis movement of torch using lead screw for welding carburetor components of sheet metal thickness of 1.2 mm. The set up with torch movement is shown in Fig. 4.

3.2 Automatic Manipulator movement

In this type of automatic setup movement of workpiece is provided by the manipulator movement and welding torch is held stationary. The stepper motor **(**Bipolar, 200 Steps/Rev, 42×38mm, 2.8V, 1.7 A/Phase) is used to drive the lead screw shaft on which the manipulator is attached. Block diagram for Stepper motor control system has been shown in Fig. 5.

The input signals receive by stepper motor is in the form of pulses. First the power supply is given to the programmable logic control then it selects the motion data signal and send it to pulse generator for controlling the motion. Then pulse generator directs the pulse signal towards motor driver which converts the pulse signal into electrical signal and send it input terminal of stepper motor. Movement of manipulator is provided by the mechanism as shown in Fig. 6.



Fig. 4 Torch mounted on trolley moving on dedicated track

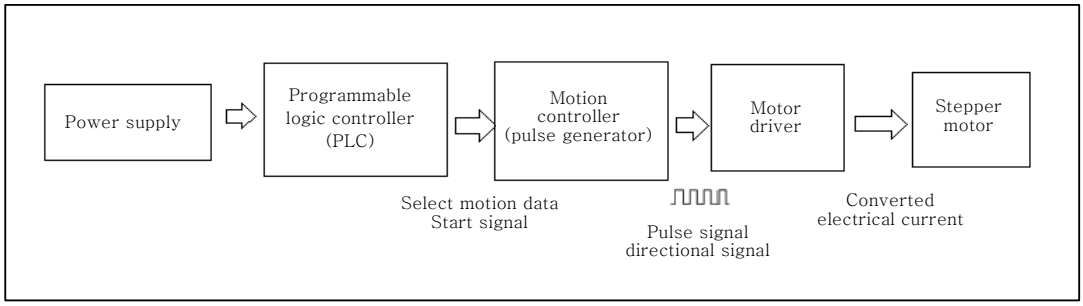


Fig. 5 Block diagram of stepper motor control system

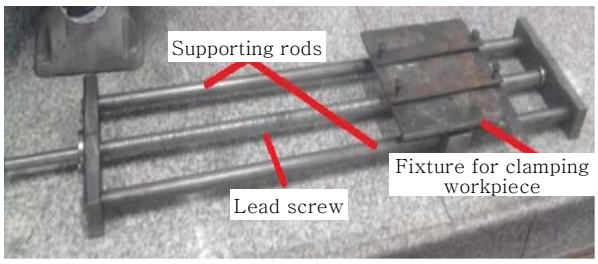


Fig. 6 Single axis driven mechanism

The wire feed system is another part of automation in pulse TIG welding system .Controlling the wire feed rate is essential parameter as it influences cooling rate of weld which effects the grain size and ultimately influences weld metal properties[14]. Welding done on thickness more than t > 6 mm requires wire feeding. When welding dissimilar metal the wire feeding becomes more important as required joint properties are obtained when the composition of filler alloy matches with the dissimilar metals to be joined.

Currently the wire feed system present in the market are MIG wire feeder systems whose cost is $ 1000 [15]. Thus it is obvious that cost incurred for manufacturing will be much higher. To decrease the cost of welding manufacture the custom made wire feed system is build.

3.3 Custom built wire feed system

The custom built wire feeder system comprises of feeder casing, feeder motor (5A, 24V), MIG Torch, Teflon liner, Copper wire Transformer (5A, 12V), Silicon bridge rectifier (406KBL), Voltage variac (4A). First the 220V, 50 Hz power supply was given to voltage variac or dimmer to vary the voltage which leads to change in fire feed rate, then output of voltage variac is given in step down transformer which decreases the voltage as per its rating. The output of decreasing voltage is then given in silicon bridge rectifier which converts ac into dc and finally the dc output is given in the dc wire feed motor. The custom build setup built along with circuit diagram is shown in Fig. 7. The total expenditure occurred was around $3000 which is much lesser as compared to feeders available in the market.

Over the years there has been lot of modifications in improving the existing TIG welding techniques. Some recent modifications in TIG welding setup in the form of variants. Some recent variants includes hot wire TIG where resistance heated filler wire is fed into the weld puddle preventing the weld pool from chilling early which allows the filler to flow out across the weld puddle resulting in smooth, attractive weld bead. Arvinda et al.[16] studied and compared the mechanical and microstructural properties of modified 9 Cr-1Mo welds produced by narrow gap hot wire and cold wire GTAW and results shows that impact properties produced at 0℃ of 9Cr-1Mo is inferior for hot wire GTAW as compared to welds produced by cold wire GTAW and at 18℃ toughness is comparable for both hot wire and cold wire GTAW.

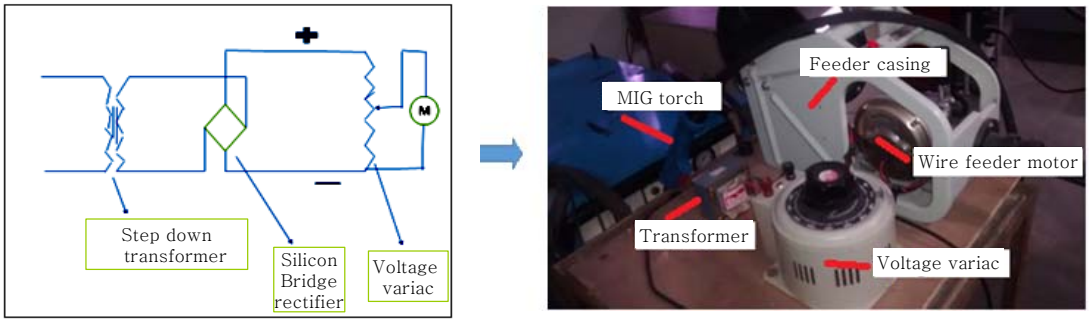


Fig. 7 Circuit diagram of custom built wire feeder and custom build wire feeder for pulse TIG

Another variant is TIG arc spot welding which is same as conventional TIG except control system which includes timing device and torch nozzle is modified. Morgan et al.[17] developed the model to predict heat transfer. The model takes into account Lorentz forces and surface tension effects. The results indicates that multiphysics simulation is useful numerical technique to study the estimated front motion and disturbance caused by sensors.

Twin electrode TIG welding is other variant of TIG. In this the arc is made to strike from two electrodes separately. The two electrodes are moved side by side along the joint to obtain a wide smooth weld without any undercuts. Zhang et al.[18] studied the physics characteristics of coupling arc formation. Arc cone shape using CCD camera was observed and arc pressure was measured. The results show that coupling arc includes two arcs that pull each other according to Lorentz force which forms one big coupling. Thus results proves that stable welding process can be achieved by twin-tungsten TIG at higher current than that of conventional TIG because of low arc pressure and high efficiency welding.

Manual programmed TIG welding is another variant where welding is achieved by use of foot controlled rheostat for starting and finishing a weld. A finger switch on TIG torch can be used to select the specific programme to achieve a desired weld. Main applications include welding tube-to-tube sheet and root pass welding in pipe work.

4. Experimental procedure

Soft aluminum alloy 5083-O and strain hardened aluminum alloy 6061-T651were used in this study. AA 5083-O Al alloy was placed in the right side and AA 6061-T651 was placed in the left side of the direction of weld as AA 6061 has higher flow strength than AA5083 aluminum alloy). A weld size of 120mm × 150mm × 6.35mm of each was used for fabrication of butt joined



(a) (b)

Fig. 8 a) Welded plate specimen of 5083-O and 6061-T651 b) Radiography

using Pulse TIG welding (Triton 220 AC/DC). Fig. 8 shows welded specimen of 5083-O and 6061-T651 and radiography. Defect free welds are obtained as shown by the radiography film

4.1 Mechanical Properties

The tensile samples were extracted from welded samples along the transverse welding direction and prepared as per ASTM E8 M04 standard as shown in fig. 9. Tensile test was carried out in 50 KN eletro-mechanical controlled universal testing machine (Tinius Olsen) with cross head speed of 2.5mm/min. The joint exhibits tensile property of 213 MPa, yield strength of 176 MPa and elongation of 12 %.

5. Conclusion

1) Automation of Pulse TIG setup is done to weld dissimilar metals. A customerized setup is build whose manufacturing cost is very low as compared to setup available in the market. This makes it a low cost auto-

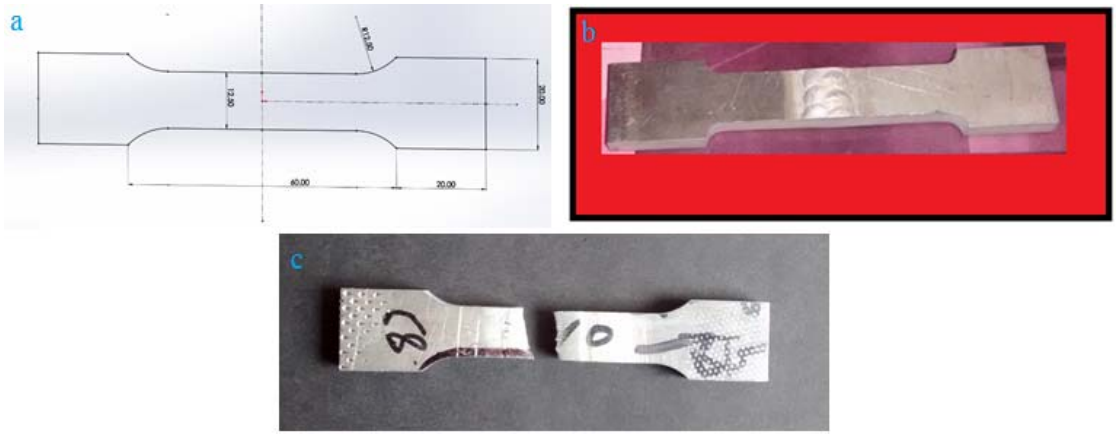


Fig. 9 Shows tensile specimen

mation of welding process.

2) Recent variants has also been discussed which are automated with few modifications compared to conventional Pulse TIG welding setups. The modifications exclusively made to enable them to be used for particular applications.

3) Defect free weld is obtained as indicated by radiography results.

4) Tensile results exhibits tensile property of 213 MPa, yield strength of 176 MPa and elongation of 12%

References

1. H.B Cary. Modern welding technology (NJ), Prentice-Hall, (1989)
2. J. Cornu. Advanced welding system, TIG and related processes. *Springer, Heidelberg* 3 (1988), 61
3. T Mohandas, R G Madhusudhana, Effect of frequency of pulsing in gas tungsten arc welding on the microstructure and mechanical properties of titanium alloy welds. *J Mater Sci Lett*, 15 (1996), 626-628
4. V Balasubramanian, V Ravisankar, RG Madhusudan, Effect of pulsed current and post weld aging treatment on tensile properties of argon arc welded high strength aluminum alloy. *Materials Science and Engineering A*, 459 (2007), 7-18
5. M Balasubramanian , V Jayabalan , V Balasubramanian, A mathematical model to predict impact toughness of pulsed current gas tungsten arc welded titanium alloy, *International Journal of Advanced Manufacturing Technology*, 35 (2008), 852-858
6. KCW Yunga, B Ralph, WB Lee, R Fenn, An investigation into welding parameters affecting the tensile properties of titanium welds, *Journal of Material Processing Technology* 63(1-3) (1997), 759-764
7. C Meran, Prediction of the optimized welding parameters for the joined brass plates using genetic algorithm. *Material Design*, 27(5) (2006), 356-363
8. Trade India catalogue, (2010)
9. BC Howard, CH Scott, Modern Welding Technology. Pearson/Prentice Hall, (2005)
10. A Karpagaraj, NS Shanmugam, K Sankaranarayanasamy, Some studies on mechanical properties and microstructural characterization of automated TIG welding of thin commercially pure titanium sheets, *Material Science & Engineering A*, 640 (2015), 180-189
11. T Otani, Titanium Welding technology, *Nippon Steel Technical Report*, 95 (2007), 88-92
12. WH Minnick, Gas Tungsten Arc Welding, Goodheart Willcox co., (1995)
13. RW Ttulanka, SS Dehankar. Automation in Sheet Metal TIG Welding Process: A Case Study, 7 (4) (2013), 3077-3083
14. Catalogue India mart, (2014)
15. AK Hussain, A Lateef, M Javed, T Pramesh, Influence of welding speed on tensile strength of welded joint in TIG welding process. *International Journal of Applied Engineering Research*, 3(1) (2010), 26-32
16. A Pai, I Sogalad, SK Albert, P Kumar, TK Mitra,S Basavarajappa, Comparison of Microstructure and Properties of Modified 9Cr-1Mo Welds Produced by Narrow Gap Hot Wire and Cold Wire Gas Tungsten Arc Welding Processes. International Conference on Advances in Manufacturing and Material Engineering AMME. *Procedia Material Science*, 5 (2014), 1482-1491
17. M Dal, PL Masson, M Carin, A model comparison to predict heat transfer during spot GTA welding. *International Journal of Thermal Sciences*, 75 (2014), 54-64
18. Z Guang-jun, L Xue-song, WU Lin, Physics characteristics of coupling arc of twin-tungsten TIG welding, 16 (2006), 813-817

Volume 4, Issue 4 (2016) 633-634 ISSN 2347 - 3258
International Journal of Advance Research and Innovation

QFD Concepts Integrated with Design of Indian Military Vehicles

MK Shukla[1], BS Chauhan[2], Ashutosh Kumar Rai[3]

[1]Department of Mechanical Engineering, Gyan Ganga College of Technology, Bargi Hills, JabalPur (MP) India
[2]Department of Mechanical and Automation Engineering, Delhi Technical Campus, K Park, Noida (UP) India
[3]Department of Mechanical Engineering, Delhi Technological University, Bawana Road, Delhi, India

Article Info

Article history:
Received 02 April 2016
Received in revised form
20 May 2016
Accepted 28 May 2016
Available online 15 December 2016

Keywords: Quality Function Deployment Product design specification, Remote controlled weapons, combat, House of quality

Abstract

The design requirement of military vehicles is very specific, unique and of high reliability. The product design is exclusively for particular purpose to suit defense requirement i.e. maneuverability, survivability and combat. The design of the product generated based on customer voice by way of Quality Function Deployment reveals optimum satisfaction of user.

The paper describes how Q.F.D. can be applied as a tool with good potential to ascertain the product design having rare specification of the product to fulfill the requirements of army personnel. QFD presents a prime facie case for being a powerful tool for the transformation of vague, imprecise, customer requirements in the clear measurable technical requirements.

1. Introduction

On focusing the country's defense scenario, it is observed that the battle zone now has been shifted from conventional to unconventional. Besides the border incursion, there is serious security problem of jungle warfare, terrorism and nexalities. Moreover to operate vehicle at difficult terrain having extreme climatic conditions is also a challenging issues while finalizing product design specification. Defense vehicle needs military solutions i.e. Armaned personal carrier, troopers, mobile command and control and home land protection. Also to meet world class and robust security system along with environments condition of the battle zone. Maneuverability, survivability, combat and road off efficacy are vital factors uncomprisely there. In addition to that interior design includes comfort, protection, fitments of communication equipment, monitoring system, electronic counter measure, manual as well as remote control weapon stations etc. are important element of design.

1.2 The Customers of Army Vehicle -

(1) Indian Military
(2) Border Security Force (BSF)
(3) Central Reserve Police Force (CRPF)
(4) Anti Terrorist Force (ATF)
(5) Indian Territory Border Police (ITBP)

1.3 Some Indian Light Armored Defense Vehicles

(1) Stallion MK S
(2) LPTA MK 13
(3) Mine Protecting Vehicle
(4) Mahindra Rak Shak
(5) Panther Model TE, 58
(6) TATA - LSV - Q DMCT

Corresponding Author,
E-mail address: mkshukla21045@gmail.com
All rights reserved: http://www.ijari.org

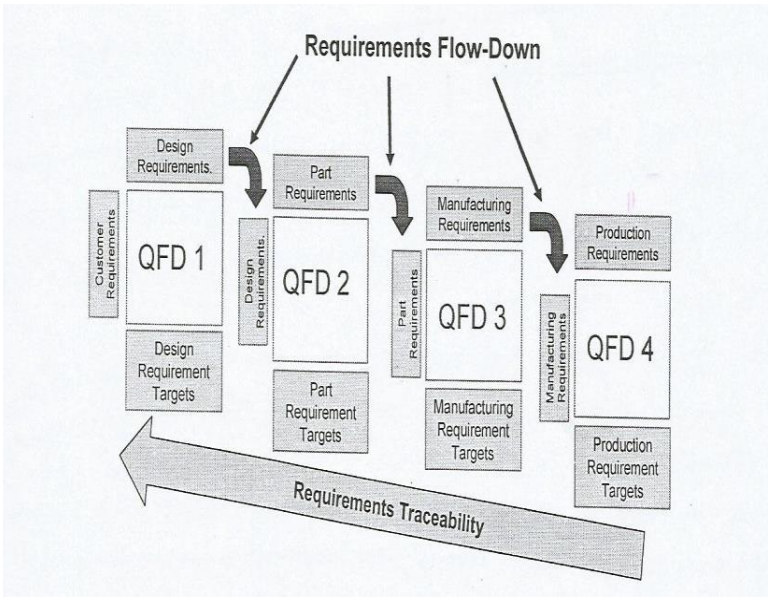


Fig.1 QFD Process

2. Review of QFD Methodology

Q.F.D. is a customer driven tool in implementing total quality management. It helps in translating customer needs into requisite technical specifications of the product at various design level then goes for its development. Q.F.D. aids in translation of vague, impressive customer needs into clear measurable technical specification. This provides a logical and systematic methodology for capturing and groaning the requirements.

QFD deals with data related to customer satisfaction identifying attribute, which should be considered while designing the product. The analysis and classification of

customer request effecting the satisfaction of customer. That is a systematic approach which can be mentioned in the following steps -

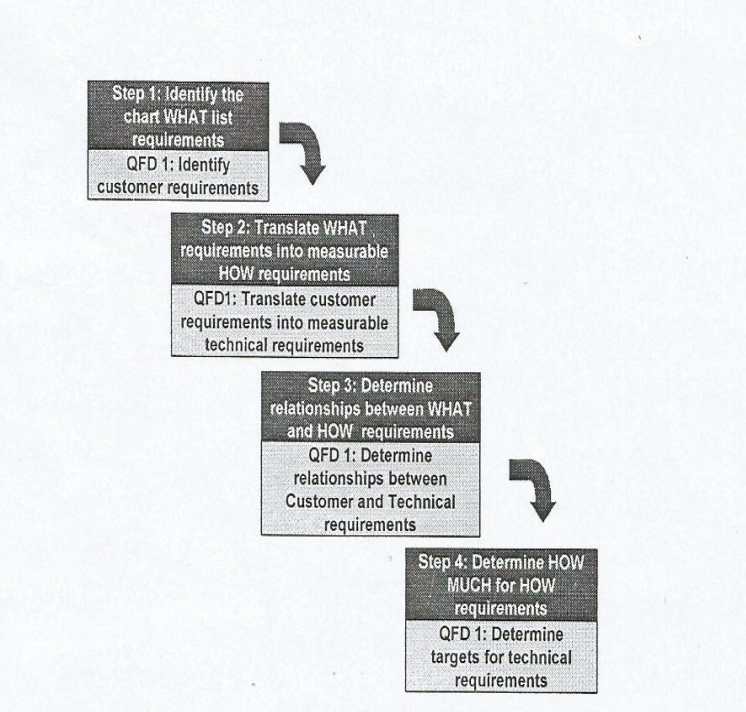


Fig. 2 Step of QFD Method

Step -1

The first step is to capture the basic needs of the customer i.e. "WHAT" is that the customer needs. This exercise requires considerable time to study and scrutinizes the requirements of the customer. Apart from that it is tactful and forceful activities to ascertain the relative importance of the requirements of the customer. The output of this process is the base of decision making while designing the product.

Step - 2

To translate voice of the customer into technical attributes, it is essential to refine the vague, ambiguous needs of the customer prior to transform into measurable technical requirements of the product.

Step - 3

As because customer need i.e. "WHAT'S" and technical requirements (How) are in good number. Therefore establishing appropriate and feasible relationship is one of the most important step of the concepts. "House of Quality" that is well known tool of Q.F.D. to capture relationship matrix is applied. In this process strong, medium and week relationship by mapping customer requirements and technical requirements are carried out.

Step - 4

House of Quality matrix provides conclusion over prioritization of customer requirement and technical requirement relationship. The result thus optimum obtained is basically applied while designing the product to achieve the satisfaction of the customer.

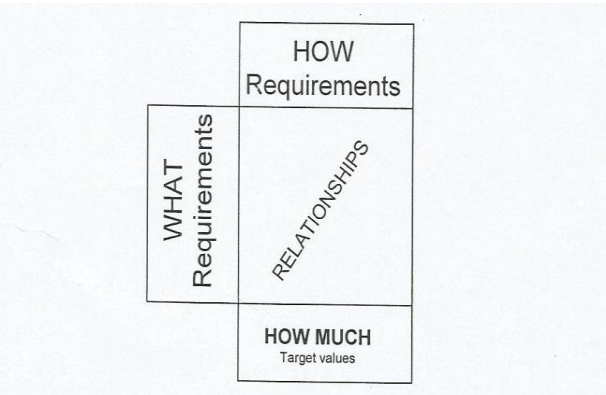


Fig.3 Basic of House of Quality

3. Concept Generation

The design of the product is undertaken based on technical specification co-related with customer attributes some optional facilities like foldable structure, first aid facilities and intercom system etc. should also be considered to delight the customer. The upgrading of the product to the tune of next generation level as well as world class version of vehicle are now expected by arms personal for quick, effective response to the attacks. Therefore product study and opinion of the customer should be studied to know the existing realities and fact about the use of the product. The customer identification and interaction with them assists to generate product specification.

4. Conclusions

Now the day's army operations are pushing from peace - keeping to irregular warfare, in border incursions, internal threat of terrorist and nexalities attack. Army needs unarmored and multirole all terrarium vehicle, specialized operation against attacks. Quality Function Deployment provides a practical and pragmatic approach that enables to enhance the quality of the product.

References

[1.] Sturat Burge. A functional approach to Quality Function Deployment System Engg. 2007

[2.] Akao Y. QFD- Integrating Customer requirement into product design. Combridge, MA: Productivity Press. 1990

[3.] ISO QFD. Portland Sept. 2010.

[4.] Clausing D, S Pugh. Enhanced Quality Function Deployment, Design & Producting Int. Honolulu 41. 1991

[5.] Midhun Eliyas IV. Hime Kiran Vithal. Concept design of Spl. Veh. Indian Military, 13, SAS Tech. Journal. 2014

[6.] John Pike. Common Remotely Operated Weapon Station, 2013.

[7.] Defense News. India needs move vehicle but spending less, 2013.

J Comput Electron
DOI 10.1007/s10825-016-0949-4

Quantum analysis based extraction of frequency dependent intrinsic and extrinsic parameters for GEWE-SiNW MOSFET

Neha Gupta[1] · **Rishu Chaujar**[1]

© Springer Science+Business Media New York 2017

Abstract This paper examines the bias-independent and bias-dependent extrinsic and intrinsic parameters of the gate electrode workfunction engineered (GEWE) silicon nanowire (SiNW) metal–oxide–semiconductor field-effect transistor (MOSFET) by considering quantum effects. The results reveal that the effect of extrinsic parameters such as the resistance, capacitance, and inductance of the electrodes is less pronounced in the GEWE-SiNW compared with the conventional SiNW or conventional MOSFET. The intrinsic transconductance of the GEWE-SiNW device can be further improved by tuning the gate metal workfunction difference, which results in shorter time constant and lower parasitic capacitance, making it suitable for radiofrequency integrated circuit (RFIC) design. It is also observed that, in the saturation region, the device exhibits improved transconductance and noticeable reduction in C_{sdx} [due to drain-induced barrier lowering (DIBL)] but the parasitic capacitance and time constant also reduce. In addition, a non-quasi-static small-signal model has been studied in terms of Z and Y parameters; the results show good agreement with the results of three-dimensional (3D) simulations at thousands of GHz.

Keywords Extrinsic and intrinsic parameters · Gate electrode workfunction engineering (GEWE) · Quantum models · Silicon nanowire (SiNW) · Z and Y parameters · Modeling

✉ Rishu Chaujar
rishu.phy@dce.edu

Neha Gupta
nehagupta@dtu.ac.in

[1] Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

1 Introduction

Accurate modeling of nonlinear transistors for RF and microwave circuits is required. As the complexity of high-frequency (HF) circuits and systems continues to increase, there is also a need to reduce design cost, enhance performance, and meet tight limits regarding spurious discharge/emission. For RF modeling, a transistor is usually divided into two parts: an intrinsic part, which includes the nonlinear elements describing the operation of the transistor, and an extrinsic part, which accounts for the metallization, substrate, and semiconductor that connect the intrinsic model to the outside world [1,2]. The intrinsic components are generally voltage dependent, whereas the extrinsic components are usually considered to be independent of voltage, i.e., passive components. As operating frequencies increase into the GHz range, the importance of the extrinsic component (the part outside the channel region) dominates over its intrinsic counterpart. Therefore, there is a need for RF models that consider the behavior of both the intrinsic and extrinsic components of a device to achieve accurate predictions from simulations of designed circuits [3,4]. Earlier work reported the effect of extrinsic and intrinsic parameters on complementary metal–oxide–semiconductor (CMOS) RF performance [5]. Moreover, Alam et al. [6] reported the extrinsic series resistance of a RF MOSFET using a small-signal model. In addition, highly scaled CMOS devices show short-channel effects (SCEs), making them inappropriate for use in HF applications [7]. To overcome this drawback, various device structures have been proposed to alleviate the issue of SCEs at high frequency [8–10]. SiNW devices are the most promising candidates due to the stronger quantum confinement, which results in higher mobility compared with bulk silicon, as well as their scalability to a few nanometers, which is challenging to obtain via conventional silicon technology [11–13].

Published online: 24 January 2017

Improved analog and RF performance of SiNW devices have already been reported in literature [14]. Also, the effect of the source-to-drain capacitance due to DIBL in the SiNW MOSFET was investigated by Cho et al. [15]. Cho et al. [16] also examined the intrinsic and extrinsic parameters of junctionless SiNW devices, including the small-signal behavior in terms of Y parameters. However, the SiNW MOSFET with a SiO_2 gate faces the major problem of lower threshold voltage than desired in the very large-scale integration (VLSI) industry. Tuning the channel doping to correct V_{th} is difficult due to the narrow cylindrical NWs, and carrier mobility is also impacted by higher channel doping. Hence, tuning via gate electrode work function engineering (GEWE) [17] represents a better solution. In this regard, Gupta et al. [18,19] demonstrated that, using GEWE on gate-all-around (GAA) SiNW devices alleviated SCEs including DIBL and hot-carrier effects (HCEs), and also showed that the GEWE-SiNW MOSFET achieved cutoff frequency of thousands of GHz, making it suitable for use in RF/microwave applications [20]. Also, the effect of extrinsic and intrinsic parameters has not been reported for the GEWE-SiNW MOSFET till now. Therefore, in this work, the effect of extrinsic and intrinsic parameters was evaluated using the pinch-off cold-FET method, then these parasitics were removed from the device under consideration, a technique known as de-embedding. Furthermore, to examine the effect of device parameters on these parameters, metal gate engineering and bias variation were also investigated. The small-signal Z and Y parameters of the GEWE-SiNW MOSFET were also studied by excluding the effect of the extrinsic parameters. This paper is an extension of [21].

2 Device description and simulation methodology

Figure 1a and b present the simulated 3D device structure of the SiNW and GEWE-SiNW MOSFET, respectively. The default structural parameters and their values for these devices are stated in Table 1. In the GEWE-SiNW MOSFET, the metal gate consists of two regions: region 1 (near the source, L1 = 15 nm) and region 2 (near the drain L2 = 15 nm), which is towards 45-nm-node CMOS technology as stated in the International Technology Roadmap for Semiconductors (ITRS) [22]. Region 1 consists of gate metal with workfunction of 4.8 eV (gold), whereas region 2 is a titanium metal gate with workfunction of 4.4 eV. Meanwhile, in the case of the SiNW MOSFET and conventional MOSFET, the entire gate metal is gold. To analyze the device performance fairly, all three devices were optimized for the same threshold voltage, i.e., 0.4 V. All simulations were carried out using the Silvaco Atlas simulator [23]. Figure 1c and d show contour plots of the electron concentration along the channel of the SiNW and GEWE-SiNW device, respectively. It is evident from these contours that, when using the GEWE scheme, the electron concentration in the channel increases, resulting in improved current driving capability compared with the SiNW MOSFET owing to the step potential profile due to the two dissimilar metal gates, thereby reducing the electric field at the drain end and thus providing hot-carrier reliability [24].

Table 1 Default design parameters of the SiNW and GEWE-SiNW MOSFET

Parameter	Value
Channel length (L_g)	30 nm
Radius of SiNW (R)	5 nm
Oxide thickness (t_{ox})	1.5 nm
Source/drain doping (N_D^+)	$5 \times 19\,\mathrm{cm}^{-3}$
Channel doping (N_A)	$1 \times 16\,\mathrm{cm}^{-3}$

The physical models invoked during simulation were electric-field-dependent and concentration-dependent carrier mobilities, Shockley–Read–Hall recombination–generation with doping-dependent carrier lifetime, inversion-layer Lombardi CVT mobility model, high-field saturation model, and quantum-mechanical model [23]. To incorporate channel quantization and ballistic transport phenomena, we adopted the quantum-mechanical model [Bohm quantum potential (BQP)], which is more precise than the drift–diffusion method in the subnanometer range, since the inversion charge layer thickness in the conducting SiNW channel is comparable to the NW dimension. The energy balance model (EBM) and hydrodynamic model (HDM) can also be used in conjunction with the BQP model, where the quantum potential adjusts the semiclassical potential in a similar way as for the continuity equations. The model introduces a position-dependent quantum potential Q, which is added to the potential energy of a given carrier type. This quantum potential is derived using the Bohm interpretation of quantum mechanics [25] and is described below:

According to Bohm's interpretation of quantum mechanics, the wave function in a polar representation can be written as

$$\psi = R \cdot \exp\left(\frac{\mathrm{j}S}{\hbar}\right), \tag{1}$$

where $\hbar$ is $h/2\pi$, R is the probability density per unit volume, and S has dimensions of action (energy $\times$ time).

Substituting Eq. (1) into the time-independent Schrödinger equation yields

$$\frac{\hbar^2}{2}\nabla\left\{M^{-1}\nabla\left[R \cdot \exp\left(\frac{\mathrm{j}S}{\hbar}\right)\right]\right\} + Q \cdot R \exp\left(\frac{\mathrm{j}S}{\hbar}\right) = E \cdot R \cdot \exp\left(\frac{\mathrm{j}S}{\hbar}\right), \tag{2}$$

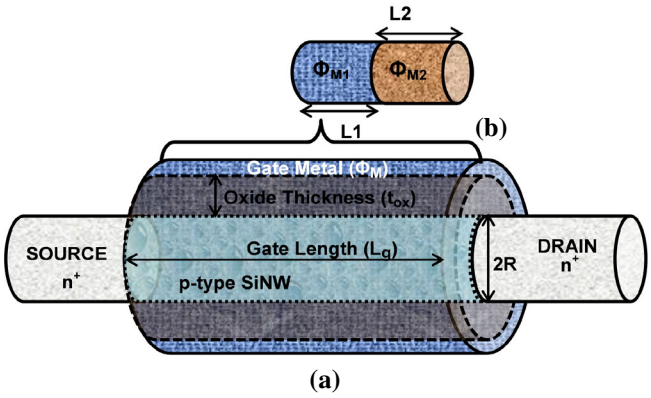


Table 1: Default Design Parameters of SiNW and GEWE-SiNW MOSFET	
Parameters	**Value**
Channel Length (Lg)	30nm
Radius of SiNW (R)	5nm
Oxide Thickness (t_{ox})	1.5nm
Source/Drain Doping (N_D^+)	5x19cm^{-3}
Channel Doping (N_A)	1x16cm^{-3}

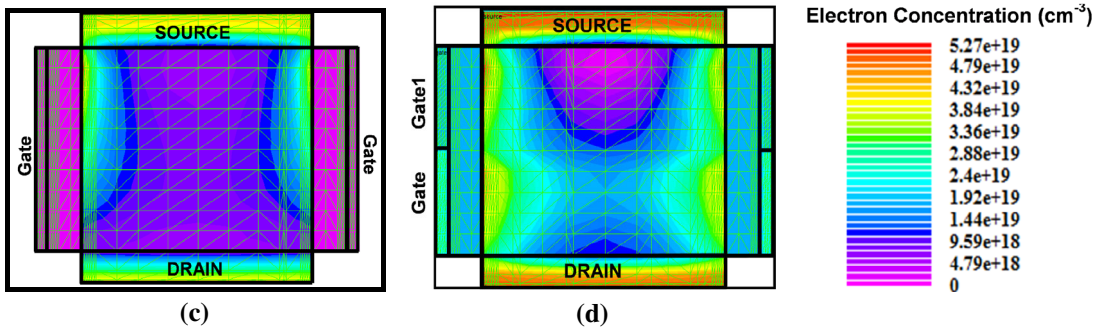


Fig. 1 3D device structure of **a** SiNW MOSFET and **b** GEWE-SiNW MOSFET, **c**, **d** two-dimensional (2D) cross-sectional contour plots of the electron concentration in the SiNW

where M^{-1} is the inverse of the effective mass tensor and $M^{-1}\nabla S$ is the local velocity of the particle associated with the wave function. The real part of Eq. 2 is interpreted as the continuity equation of the probability density, while the imaginary part states that the total energy E is conserved and equal to the sum of the potential (Q) and kinetic energy ($0.5\,M^{-1}\nabla S$). Thus, the quantum potential is defined as

$$Q = \frac{-\hbar^2}{2}\frac{\nabla\left(M^{-1}\nabla R\right)}{R}. \tag{3}$$

The potential Q is obtained from the single-particle Schrödinger equation using the mean-field approximation and considering an effective quantum Bohm potential defined as the weighted average potential for all carriers constrained in the confining potential. Therefore, the effective quantum potential can be written as

$$Q_{\text{eff}} = \frac{-\hbar^2}{2}\gamma\nabla\frac{[M^{-1}\nabla\,(n^{\alpha})]}{n^{\alpha}}, \tag{4}$$

where α and γ are two adjustable parameters, M^{-1} is the inverse effective mass tensor, n is the electron (or hole) density, and h is Planck's constant.

Additionally, numerical techniques such as the biconjugate gradient stabilized (BICGST) method were considered to obtain solutions with improved convergence for the 3D device structures [23]. Default simulator coefficients were employed for all parameters. To calibrate the BQP model, the two adjustable parameters were tuned to $\alpha = 0.5$ and $\gamma = 1.2$ to obtain the best fit with experimental data. Calibration of the model parameters used in the simulation was performed according to the experimental results [26] using the above-mentioned models. Figure 2 shows the simulated I_d–V_{gs} characteristic of the 30-nm cylindrical gate SiNW MOSFET at $V_{ds} = 1.0$ V and the transfer characteristic of extracted data from [26]. The results are in close proximity, thus validating the choice of model parameters used in the simulation.

2.1 Fabrication feasibility

The fabrication feasibility of the SiNW MOSFET has been reported in literature using several integration schemes. Yang et al. [27] fabricated a vertical gate-all-around silicon nanowire MOSFET. Rustagi et al. [28] reported fabrication of CMOS inverter-based GAA SiNW MOSFETs using a top-down approach. Moreover, to realize the GEWE architecture, numerous schemes have been suggested, such as tilt angle evaporation metal gate deposition [17], a metal interdiffusion process [29], and the fully silicide (FUSI) metal gate technique [30]. Furthermore, CMOS with dual metal gate has also been successfully fabricated with poly-Si gate doping control of the source and drain side gate individually [31]. Thus, the

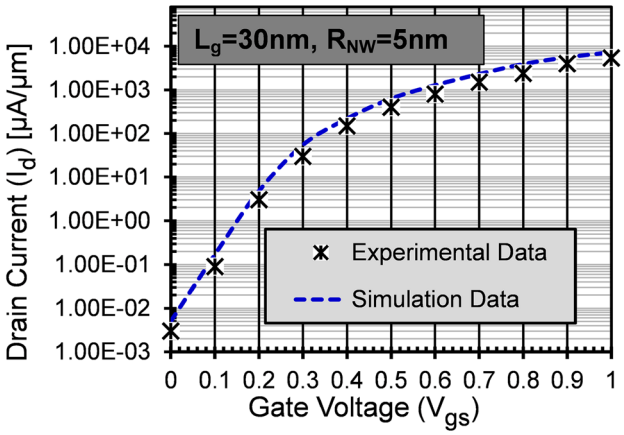


Fig. 2 Calibration with published result: I_{ds}–V_{gs} characteristic of the silicon nanowire MOSFET at $V_{ds} = 1.0$ V [26]

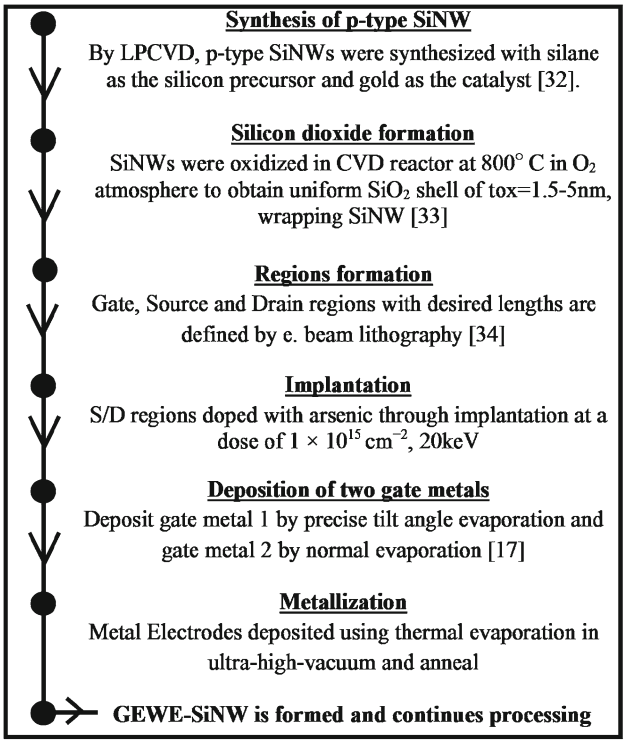


Fig. 3 Summary of typical process flow of a GEWE-SiNW MOSFET [17,32–34]

proposed device (GEWE-SiNW MOSFET) can be fabricated using the above design schemes, in which the advantages of the GEWE scheme are combined with the SiNW MOSFET structure, yielding a promising design for system-on-chip and RF/microwave applications. A proposed summary of the process flow, outlining the fabrication method of the GEWE-SiNW MOSFET and its integration with the standard CMOS process, is shown in Fig. 3.

3 Results and discussion

3.1 Extrinsic parameters

When a MOSFET is operated at low frequency, the extrinsic parameters behave as virtual components in the I–V expression to account for the direct-current (DC) voltage drop across these resistances, an effect obscured for high-frequency signals. Therefore, external mechanisms for these components must be included beyond the intrinsic model to precisely describe high-frequency device behavior [35]. Furthermore, the extrinsic network plays a significant role in determining the scaling rules for the complete transistor model. The extrinsic network has to describe the electrical behavior of the device layout, and the way in which this layout changes with device size determines how the extrinsic network component values scale. However, detailed knowledge of the extrinsic parameters is also essential for precise modeling of RF transistors, so that the extrinsic components can be de-embedded accurately to gain access to the intrinsic behavior of the transistor. To extract the extrinsic parameters, the MOSFET is biased at zero gate and drain bias, which is known as the zero-cold FET condition [36,37]. The process of extracting the parameters is described in Fig. 4a. C_{gdo} and C_{gso} are the extrinsic gate-to-drain and gate-to-source capacitance, respectively. R_{SO} and R_{DO} are the source and drain extrinsic resistance, respectively. L_{SO}, L_{DO}, and L_{GO} are the source, drain, and gate extrinsic inductance, respectively, as shown in the equivalent circuit in Fig. 4b.

In this subsection, the extrinsic parameters of the GEWE-SiNW MOSFET are examined as functions of frequency, then compared with conventional devices having identical geometry. As is evident from Fig. 5a, with increase in the frequency, all parasitic resistances due to the gate, source, and drain remain almost constant. Moreover, in the SiNW, their effects are comparatively smaller compared with in the conventional MOSFET, and with incorporation of the GEWE scheme, the resistance effect reduces appreciably. Similarly, both extrinsic capacitances and inductances also reduce in our proposed device compared with the SiNW and conventional MOSFET, as shown in Fig. 5b, c.

3.2 Intrinsic parameters

Direct techniques for extracting intrinsic small-signal parameters were performed after de-embedding the effect of extrinsic parameters such as C_{gdo}, C_{gso}, R_{SO}, R_{DO}, L_{SO}, and L_{DO}. Figure 6 shows the non-quasi-static equivalent circuit of the MOSFET used to extract the intrinsic parameters of the GEWE-SiNW, conventional silicon nanowire, and conventional MOSFET, which is based on conventional MOSFET small-signal microwave modeling [38]. Here, R_{GS} and R_{GD} represent the distributed channel resistance, and C_{DS} and

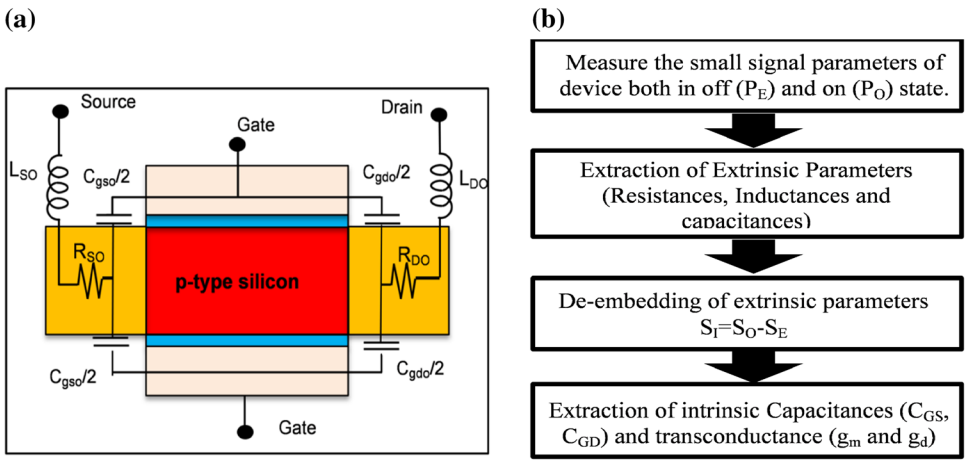


Fig. 4 **a** Procedure for extracting extrinsic and intrinsic parameters. **b** Schematic representation of equivalent circuit of SiNW MOSFET in cold FET pinch-off condition ($V_{gs} = 0\,V$, $V_{ds} = 0\,V$)

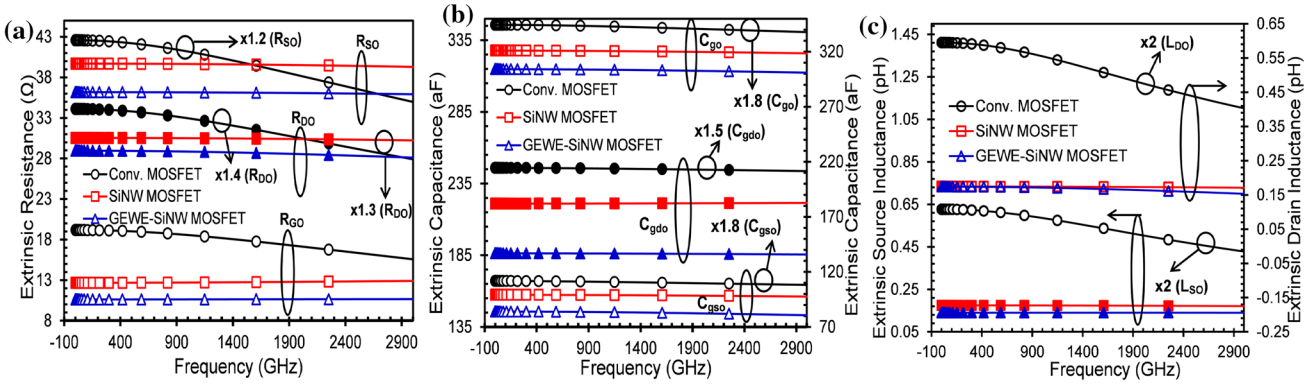


Fig. 5 **a** Extrinsic resistance, **b** capacitance, and **c** inductance as functions of frequency for the conventional, SiNW, and GEWE-SiNW MOSFET

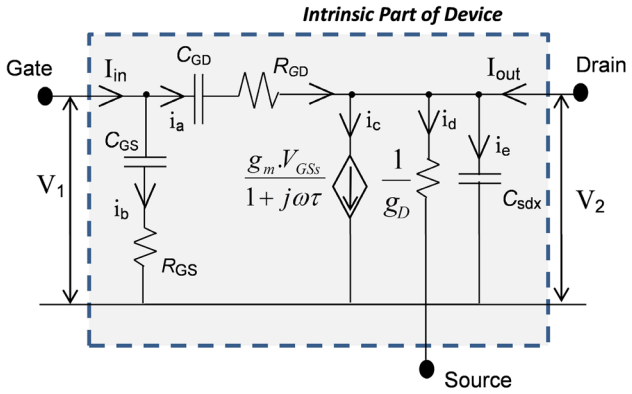


Fig. 6 Non-quasi-static (NQS) equivalent circuit model of the SiNW MOSFET in RF regime operating in strong inversion region

C_{GS} are the intrinsic gate-to-drain and gate-to-source capacitance, respectively. g_D and g_m are the intrinsic drain and gate transconductance, respectively.

Figure 7a illustrates the bias dependence of the small-signal parameters in terms of the gate and drain transconductance of the GEWE-SiNW and conventional SiNW MOSFET at gate and drain bias, respectively. As is evident from Fig. 7a, there is a significant increment in g_m with an increase in V_{gs}, compared with the SiNW MOSFET. This is due to the enhanced current driving capability of the GEWE-SiNW MOSFET compared with the conventional SiNW MOSFET, thereby improving the gain of the device. Furthermore, g_D decreases with increase in V_{ds}, but its value is comparatively higher in the case of the GEWE-SiNW MOSFET, since g_D is inversely related to the output resistance. It is evident that, when the device is in the active region ($V_d = 0$–$0.3\,V$), g_D decreases rapidly, indicating that the output resistance decreases with enhancement of the on-current; while in the saturation region ($V_d = 0.4$–$1.0\,V$), g_D becomes almost constant, which results in saturation of the drain current, as shown in Fig. 7a. Parasitic capacitances such as C_{GS}, C_{GD}, and C_{GG} in RF amplifiers degrade the performance in terms of low gain and, in some cases, may cause such amplifiers to oscillate. In digital logic circuits, the rise and fall times of the digital signal greatly influence the maximum achievable device speed [39]. Thus, it is necessary to examine the effect of parasitic capacitances at HF. The capacitances of the SiNW MOSFET turned out to be smaller than those of the conventional MOSFET, as shown in Fig. 7b. On incorporation of the GEWE scheme into the SiNW MOSFET, the parasitic capacitances decrease further with increase in fre-

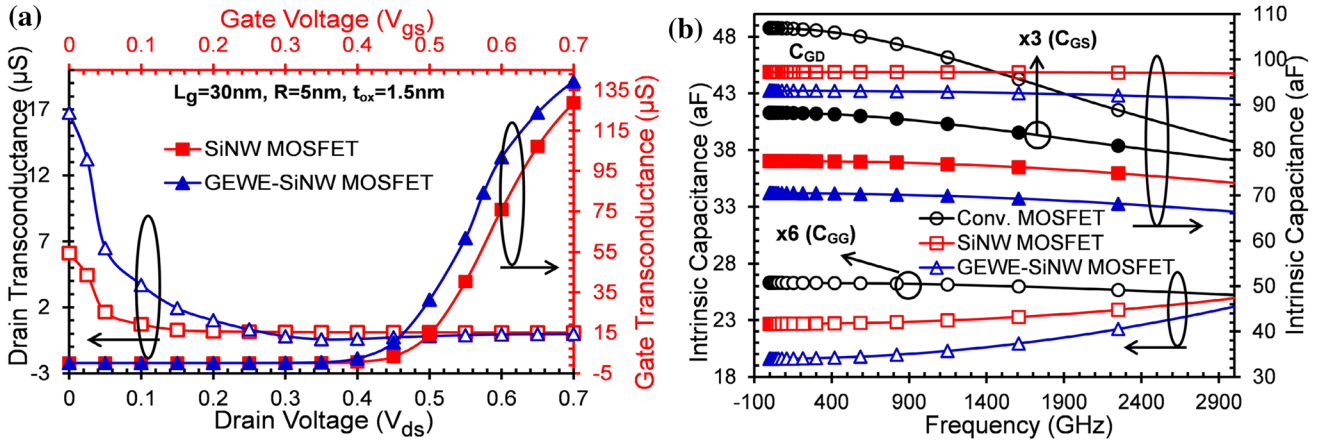


Fig. 7 **a** Intrinsic gate and drain transconductance of a conventional, SiNW, and GEWE-SiNW MOSFET as functions of bias voltage, **b** intrinsic capacitance as function of frequency for conventional, SiNW, and GEWE-SiNW MOSFET

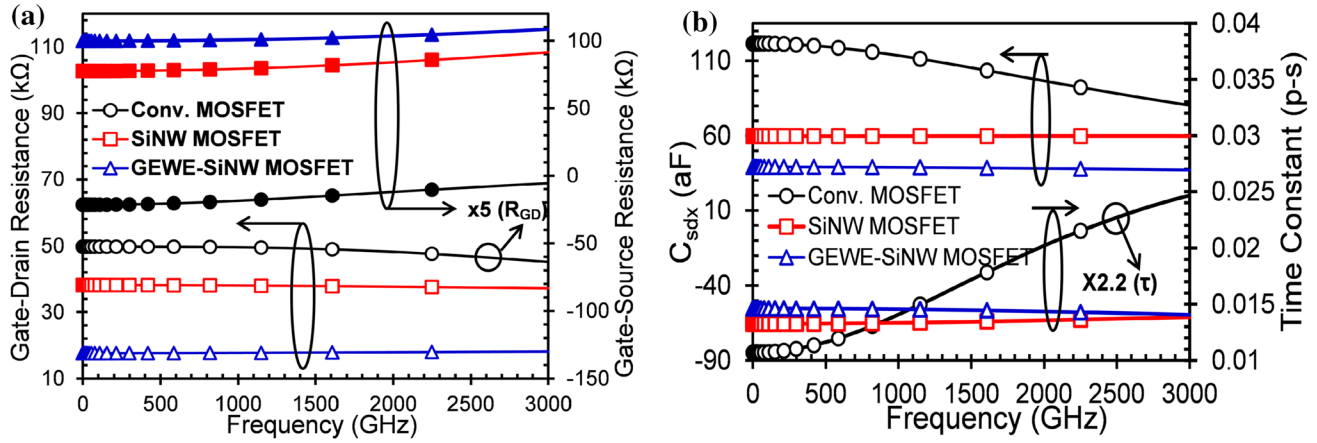


Fig. 8 **a** Intrinsic gate-to-drain and gate-to-source resistances of conventional, SiNW, and GEWE-SiNW MOSFET as functions of frequency, **b** C_{sdx} and time constant of a conventional, SiNW, and GEWE-SiNW MOSFET as functions of frequency

quency. This is due to improved screening of the channel region from drain-bias variations. Hence, lower values of the intrinsic parasitic capacitances mean lower intrinsic delay, which is suitable for switching applications such as in logic gates. At low frequency, the gate of the MOSFET is treated as purely capacitive. However, this assumption fails when the frequency exceeds the GHz range [40]. Also, the distributed gate resistance plays a major role in evaluating the noise figure and power gain of the transistor. Figure 8a exhibits the distributed channel resistances, i.e., R_{GD} and R_{GS}, for the GEWE-SiNW, SiNW, and conventional MOSFET as functions of frequency. R_{GD} for the GEWE-SiNW MOSFET is much smaller than for its conventional counterparts, resulting in the maximum power gain and maximum oscillation frequency.

Also, in highly scaled MOSFETs, the unexpected change of the inversion channel charge is due to the DIBL effect. Since the source and drain capacitance of the transistor are determined by the inversion channel charge, the value of C_{SD} is affected by DIBL [41]. To study the effect of DIBL on C_{SD} in detail, one can use C_{sdx} as an additional component reflecting the charge variation due to DIBL in short-channel MOSFETs [15]. Figure 8b shows the variation of C_{sdx} as a function of frequency. The effect of DIBL is highly pronounced in the conventional MOSFET at HF compared with the SiNW MOSFET. With incorporation of the GEWE, SCEs such as hot-carrier effects and DIBL are reduced, as reported in literature [18]. It is clearly shown that the DIBL effect is reduced in the GEWE-SiNW MOSFET due to the reduced C_{sdx} owing to a decline in the electric field at the drain end because of the lower metal workfunction. Figure 7b shows the time constant of the conventional, SiNW, and GEWE-SiNW MOSFET as a function of frequency. The time constants for charging delay are obtained as $R_{GS}C_{GS}$ and $R_{GD}C_{GD}$, and the charging delay is affected by the transport delay τ [38]. It is evident from Fig. 7b that τ for the SiNW MOSFET is greater than for the conventional MOSFET due to the higher inversion channel charge owing to its cylindrical geometry.

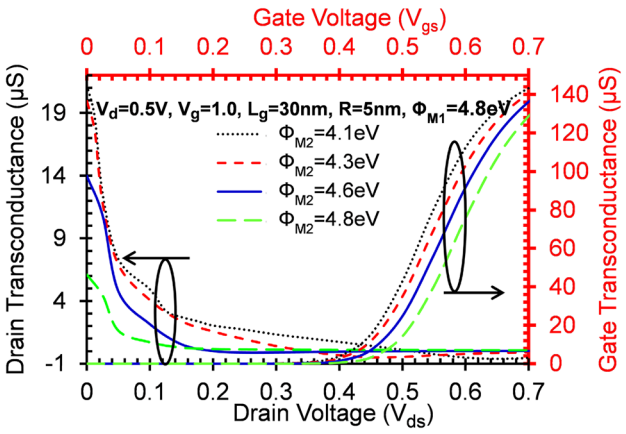


Fig. 9 Intrinsic gate and drain transconductance of the GEWE-SiNW for different metal workfunctions

On incorporation of two dissimilar metal gates, the time constant increases. Therefore, the distributed channel resistance R_{GS} of the GEWE-SiNW MOSFET is higher than that of the conventional device.

3.2.1 Effect of gate metal workfunction engineering

In this subsection, the effect of metal gate engineering on the intrinsic parameters of the GEWE-SiNW MOSFET is investigated by keeping the workfunction at the source side constant (4.8 eV) while varying the workfunction of the metal gate at the drain end. Figure 9 shows the variation of the bias-dependent gate and drain transconductance for different workfunctions at the drain end. With increase in the workfunction difference, the gate transconductance (g_m) increases due to improved carrier transport efficiency and enhanced current driving capability, hence delivering higher gain. The drain transconductance (g_d) decreases with increase in the drain bias, an effect that becomes more pronounced as the workfunction difference increases, as shown in Fig. 9.

Moreover, with increase in the workfunction difference, there is a prominent reduction in C_{GS} and C_{GD} with increase in frequency, as observed in Fig. 10a, b, due to enhanced screening of the potential from bias variations, thus reducing the turn-on delay time. Similarly, with increase in the workfunction difference, the distributed R_{GD} decreases and R_{GS} increases, and these gate resistances are independent of frequency, as shown in Fig. 11a. It is evident from Fig. 11b that, by decreasing the workfunction of the metal gate at the drain end, the effect of C_{sdx} reduces in the GHz frequency regime, directly indicating that a higher workfunction difference leads to immunity against DIBL and hence SCEs. Furthermore, the time constant is also enhanced by 32 % with an increase in the workfunction difference, as shown in Fig. 11b, due to the enhancement in the on-current and carrier mobility.

3.2.2 Bias-dependent intrinsic parameters

In the previous section, the intrinsic parameters of the GEWE-SiNW MOSFET were evaluated as functions of frequency; the results showed that these parameters are almost constant with frequency in the GHz range. In this section, the effect of bias variation on the intrinsic parasitic capacitance, transconductance, distributed gate capacitance, and time constant is investigated. The MOSFET gate transconductance ($g_m = \partial I_{DS}/\partial V_{GS}$) is defined in the linear regime and used as an indirect monitor of the inversion charge carrier mobility. The transconductance is an important parameter needed for designing high-frequency amplifiers. As is evident from Fig. 12a, g_m increases remarkably as the drain voltage is increased from 0.1 to 0.5 V. Figure 12a also indicates the dependence of the drain transconductance on the drain voltage for different gate voltages. As the drain voltage is increased, the transconductance increases and then reduces to zero in the saturation region. The rate of decrease of the

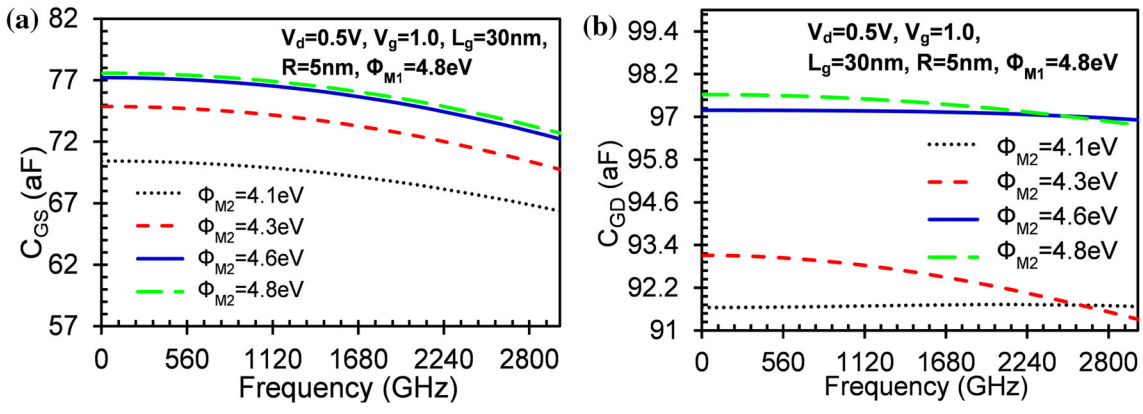


Fig. 10 Intrinsic capacitance of the GEWE-SiNW MOSFET as function of frequency for different gate metal workfunctions

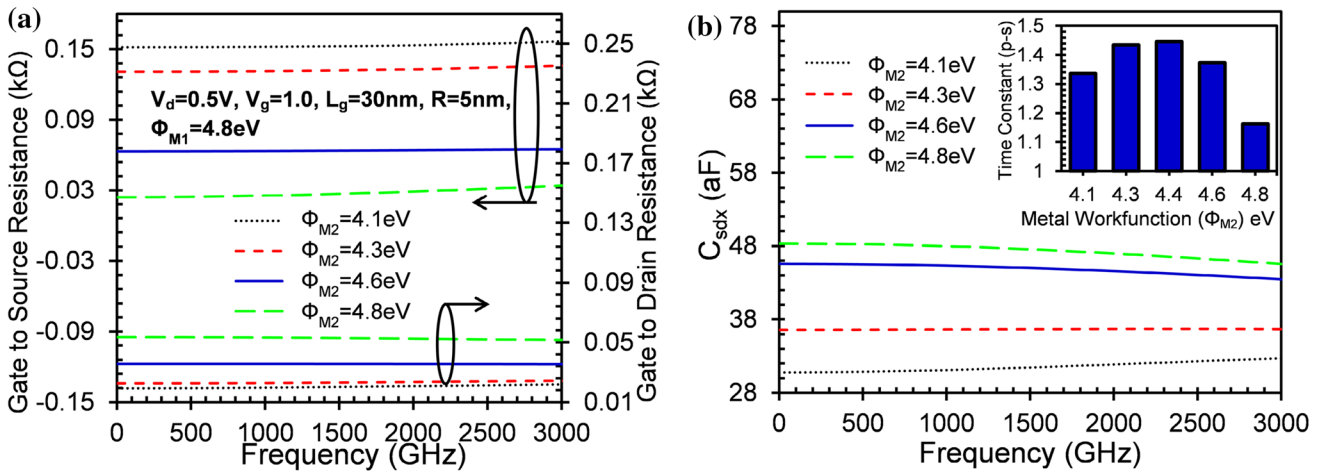


Fig. 11 **a** Intrinsic gate-to-drain and gate-to-source resistance of the GEWE-SiNW MOSFET for different metal workfunctions, **b** C_{sdx} and time constant of the GEWE-SiNW MOSFET for different metal workfunctions

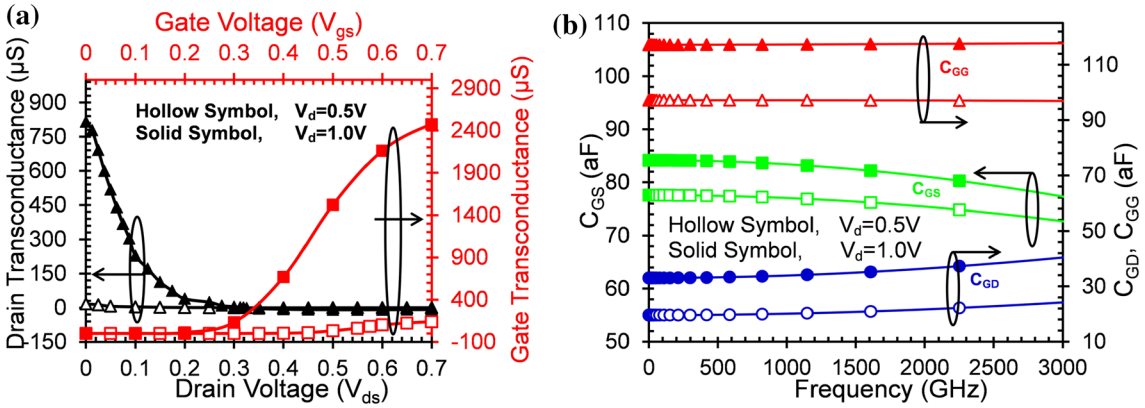


Fig. 12 **a** Intrinsic gate and drain transconductance of the GEWE-SiNW MOSFET for different bias values, **b** intrinsic bias-dependent parasitic capacitance of the GEWE-SiNW MOSFET

drain transconductance with the drain voltage is greater than that with the gate voltage.

The effect of the drain bias on the parasitic capacitance is shown in Fig. 12b. It is evident from this figure that, with increase in the drain bias, all the parasitic capacitances increase due to the shift of the pinch-off point towards the source side, which leads to accumulation of charge carriers, resulting in an increment in C_{GS}. To maintain charge neutrality, charges near the drain decrease, hence C_{GD} decreases, as shown in Fig. 12b. Figure 13a indicates the variation of C_{sdx} as a function of frequency for different drain bias values. It is evident that, with increase in the drain bias, the DIBL effect is reduced, as reflected in C_{sdx}. This reduction is due to the improved gate controllability of the GEWE-SiNW MOSFET over the channel due to the cylindrical gate and step in the conduction-band energy profile in the channel region, which itself is a result of the metal gate workfunction difference. Also, both distributed gate resistances (R_{GS} and R_H) decrease as the device moves from the linear to saturation region, as is evident from Fig. 13a. Moreover, the time constant of the GEWE-SiNW MOSFET degrades with increase in the drain bias due to enhancement of the inversion charge carriers, as shown in Fig. 13b.

3.3 Small-signal model parameters: *Z* and *Y*

The small-signal behavior of a MOSFET is generally expressed in terms of the impedance (Z parameters) and admittance (Y parameters), being commonly employed for networks that operate at RF/microwave frequencies. Based on knowledge of the parasitic parameters, the intrinsic Z parameters can be determined in a similar way as the intrinsic parameters are extracted. The choice of the parameter used is generally determined by suitability or usage; e.g., if shunt feedback is applied in RF amplifiers, Y parameters are most convenient, whereas Z parameters are recommended for series feedback. After de-embedding the extrinsic parameters, the intrinsic Z and Y parameters were evaluated by solving the small-signal equivalent model as shown in Fig. 6.

The total current flowing at the input port (gate) is $i_a + i_b$.

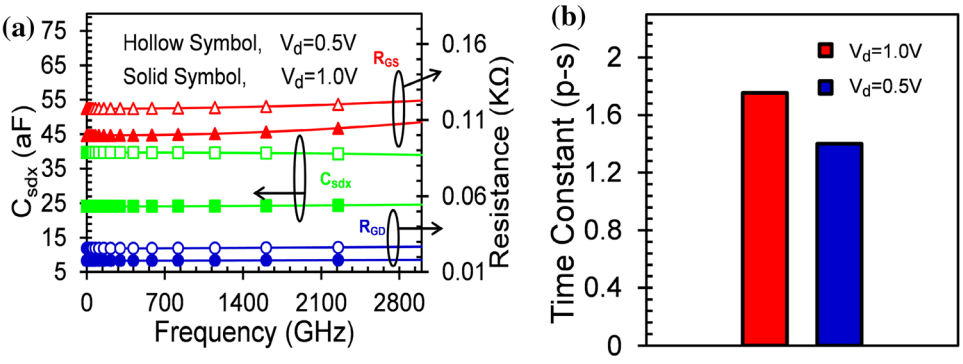


Fig. 13 C_{sdx}, distributed gate resistance, and time constant of the GEWE-SiNW MOSFET for different drain bias values

The total current flowing at the output port (drain) is $i_c + i_d + i_e - i_b$.

Applying nodal analysis,

$$\frac{V_1}{X_1} + \frac{V_1 - V_2}{X_2} = 0 \quad \text{[at input node]}, \tag{5}$$

$$\frac{V_2}{X_3} + \frac{V_2 - V_1}{X_2} + \frac{g_m V_{gs}}{1 + j\omega\tau} = 0 \quad \text{[at output node]}, \tag{6}$$

where $X_1 = \frac{1+j\omega R_{GS}C_{GS}}{j\omega C_{GS}}$, $X_2 = \frac{1+j\omega R_{DS}C_{DS}}{j\omega C_{DS}}$, and $X_3 = \frac{1}{g_D + j\omega C_{sdx}}$.

Solving Eqs. 1 and 2 using the approximations $\omega^2 R_{GS}^2 C_{GS}^2 \ll 1$, $\omega^2 R_{GD}^2 C_{GD}^2 \ll 1$, and $\omega^2\tau^2 \ll 1$ yields

1. The driving-point short-circuit input admittance

$$Y_{11} \cong \omega^2(a_1 + a_2) + j\omega(C_{GS} + C_{GD}) \tag{7}$$

2. The driving-point short-circuit output admittance

$$Y_{22} \cong g_D + \omega^2 a_2 + j\omega(C_{GD} + C_{sdx}) \tag{8}$$

3. The reverse transfer admittance

$$Y_{12} \cong -\omega^2 a_2 - j\omega C_{GD} \tag{9}$$

4. The forward transfer admittance

$$Y_{21} \cong -\omega^2 a_2 - j\omega C_{GD} + g_m(1 + \tau) \tag{10}$$

Similarly, the Z parameters can be evaluated from the Y parameters as follows:

1. The driving-point input impedance

$$Z_{11} \cong \frac{g_D + \omega^2 a_2 + j\omega(C_{GD} + C_{sdx})}{\xi} \tag{11}$$

2. The open-circuit reverse transfer impedance

$$Z_{12} \cong \frac{\omega a_2 + j\omega C_{GD}}{\xi} \tag{12}$$

3. The open-circuit forward transfer impedance

$$Z_{21} \cong \frac{\omega^2 a_2 + j\omega C_{GD} - g_m(1 + \tau)}{\xi} \tag{13}$$

4. The driving-point output impedance

$$Z_{22} \cong \frac{\omega^2(a_1 + a_2) + j\omega(C_{GS} + C_{GD})}{\xi} \tag{14}$$

where $\xi = [j\omega(C_{GS}g_D + C_{GD}g_D + C_{GD}g_m + \tau C_{GD}g_m) + \omega^2(a_1 g_D + a_2 g_D - C_{GS}C_{sdx} - C_{GD}C_{sdx} + a_2 g_m(1 + \tau) + C_{GD}^2) + j\omega^3(a_1 C_{GD} + a_2 C_{GS} + C_{sdx}(a_1 + a_2)) + a_1 a_2 \omega^4]$.

$a_1 = R_{GS}C_{GS}^2$, $a_2 = R_{GD}C_{GD}^2$.

Z_{11} and Z_{22} are essentially used for impedance matching in high-frequency applications. For impedance matching, the input impedance should be equal to the output impedance, which is necessary for maximum power transfer from source to load. Figure 14a, b shows the variation of the real component of the input and output impedance of the conventional, SiNW, and GEWE-SiNW MOSFET as functions of frequency. It is clearly shown that the input impedance of the GEWE-SiNW device is higher than that of its conventional counterparts owing to the cylindrical geometry and GEWE scheme. Furthermore, when tuning the metal workfunction at the drain end, both the input and output impedance improve as the frequency increases in the GHz range, as shown in Fig. 14c, d. For RF amplifiers, ideally Z_{12} should be as high as possible so that maximum impedance can transfer from input to output. It is clear from these results that Z_{12} for the SiNW MOSFET is greater than for the conventional MOSFET due to the reduced SCEs and maximum volume inversion, and with the incorporation of GEWE engineering into the SiNW device, Z_{12} improves, as shown in Fig. 14d. Furthermore, with increasing gate workfunction difference in the GEWE-SiNW device, Z_{12} is enhanced due to the enhancement in

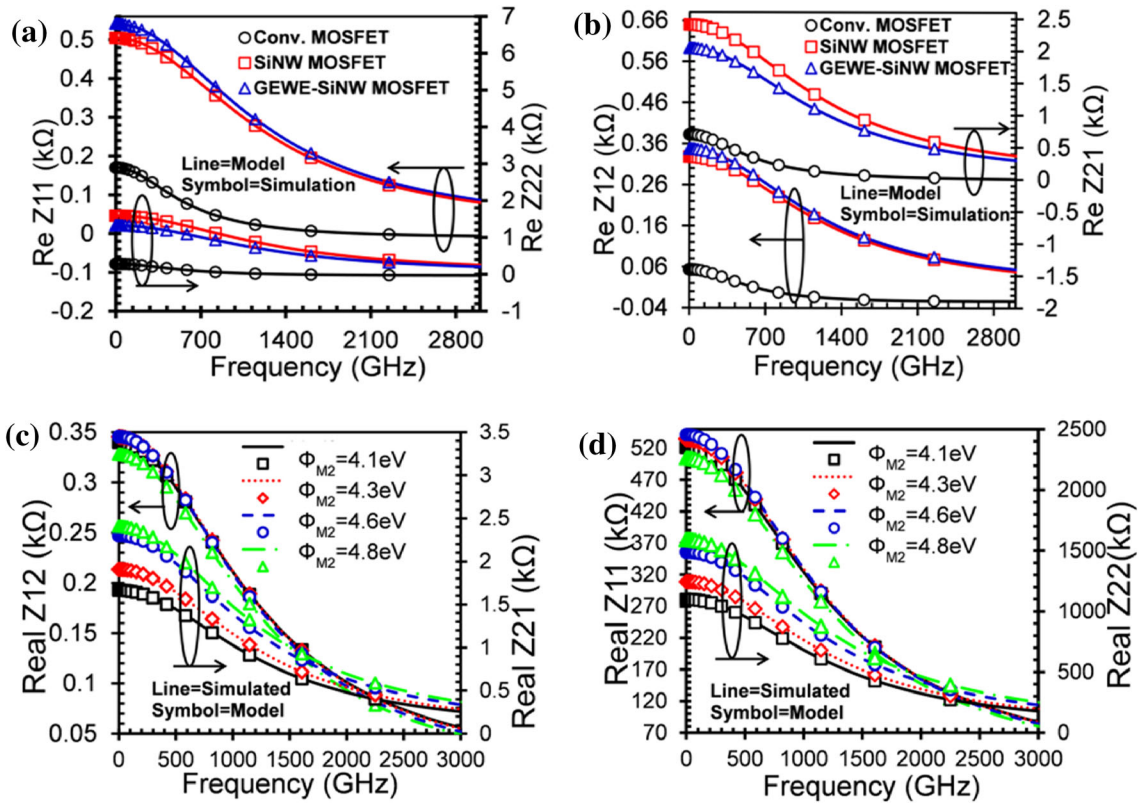
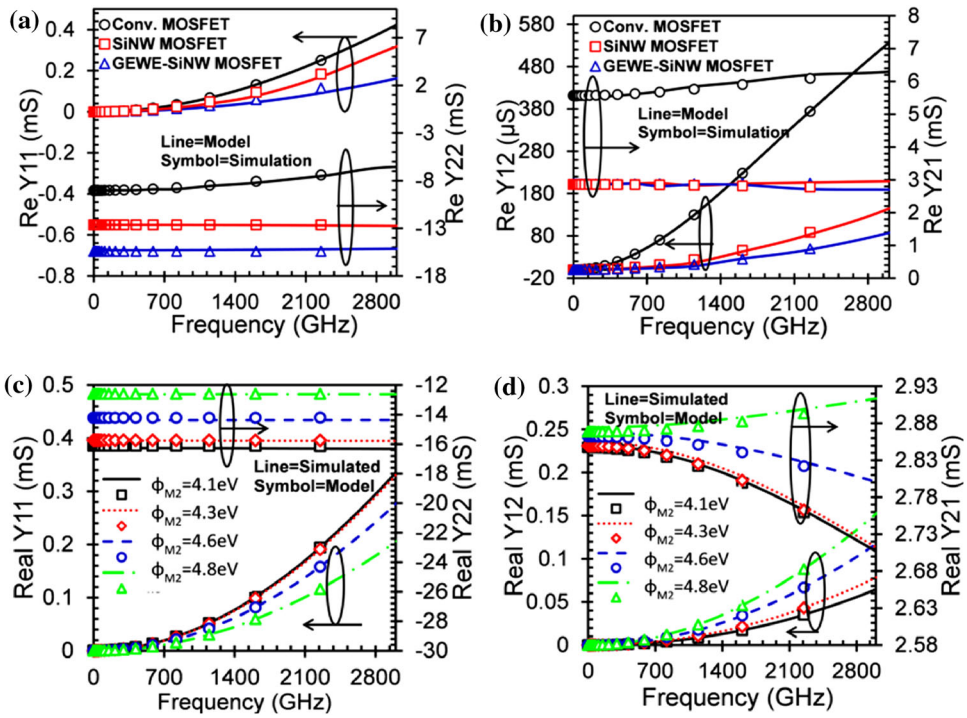


Fig. 14 **a**, **b** *Z* parameters of conventional, SiNW, and GEWE-SiNW MOSFET; **c**, **d** *Y* parameters of GEWE-SiNW MOSFET for different gate metal workfunctions

Fig. 15 **a**, **b** *Y* parameters of conventional, SiNW, and GEWE-SiNW MOSFET; **c**, **d** *Y* parameters of GEWE-SiNW MOSFET for different gate metal workfunctions



the carrier mobility and hence transport efficiency, as is evident from Fig. 14d. At low frequency, alternating-current (AC) voltages and currents can be easily measured, but at HF, measurement of these parameters becomes exceptionally complicated. Another set of parameters, i.e., the *S* parameters, can be measured at high frequency.

Table 2 Extracted parameters for all three devices for the linear and saturation regions

Parameter	Device					
	Conventional MOSFET		SiNW MOSFET		GEWE-SiNW MOSFET	
	$V_d = 0.5\,V$	$V_d = 1.0\,V$	$V_d = 0.5\,V$	$V_d = 1.0\,V$	$V_d = 0.5\,V$	$V_d = 1.0\,V$
C_{cs}(aF)	251.5	287.87	76.2	81.32	69.25	82.189
C_{GD}(aF)	310.4	340.87	97.2	103.34	92.54	117.44
R_{GS}(kΩ)	14.88	12.34	81.8	79.34	118.9	102.19
R_{GD}(kΩ)	244.65	220.98	37.8	31.98	26.1	17.82
g_m(μS)	50	65.96	128	200.98	139.3	2470
g_D(μS)	0.1	1.02	0 711	7.47	3.70	228.3
τ(ps)	0.431	0.65	1.161	1.4	1.39	1.89
C_{sdx}(aF)	309	289.34	59.829	52.56	38.57	24.31
R_{SO}(Ω)	46.8	46.8	39.585	39.585	36 11	36.11
R_{DO}(Ω)	44.14	44.14	39.585	39.585	28.68	28.68
L_{SO}(pH)	1.07	1.07	0.1725	0.1725	0.1399	0.1399
L_{DO}(pH)	1.014	1.014	0.1735	0.1735	0.167	0.167
C_{GSO}(aF)	298	298	157	157	144	144
C_{GDO}(aF)	319	319	182	182	136	136

Table 3 Extracted parameters for the GEWE-SiNW MOSFET for different metal workfunctions (Φ_{M2})

Parameter	Workfunction (Φ_{M2}) GEWE-SiNW MOSFET ($\Phi_{MI} = 4.8\,eV$, $V_d = 0.5\,V$, $V_g = 1.0\,V$)			
	4.1	4.3	4.6	4.8
C_{GS}(aF)	69.25	73.4	75.80	76.2
C_{GD}(aF)	91.7	92.81	97.14	97.6
R_{GS}(kΩ)	152	132.2	63.87	27.38
R_{GS}(kΩ)	20.5	23.16	35.53	52.78
g_m(μS)	143.7	140	136	128
g_D(μS)	4.92	4.07	2.146	0.711
τ(ps)	1.54	1.478	1.378	1.161
C_{sdx}(aF)	31.53	36.72	44.93	46.67

However, as the Y parameters are more closely related to the device physics and design, the S parameters are converted to Y parameters. Hence, study of the Y parameters becomes necessary. Figure 15a shows that the real component of the input admittance of the GEWE-SiNW MOSFET is comparatively lower than for its conventional counterparts, and its value gradually increases with increase in frequency. Similarly, Y_{22} is higher for the GEWE-SiNW device compared with the SiNW or conventional MOSFET, indicating that the current driving capability of the GEWE-SiNW MOSFET is enhanced due to the metal workfunction engineering, which further enhances the carrier velocity and hence on-current, as shown in Fig. 15a. The Y_{12} and Y_{21} values for the GEWE-SiNW MOSFET are lower and higher, respectively, as shown in Fig. 15b. Y_{21} is defined as the forward transconductance of the device; ideally, its value should be as high as possible to achieve high-gain RF amplifiers.

Also, for a workfunction difference of 0.7 eV, the small-signal performance of the device in terms of the real component of Y_{11}, Y_{12}, Y_{21}, and Y_{22} improves, as shown in Fig. 15c, d, owing to the greater enhancement in the transconductance. The modeled Z and Y parameters (lines) generated using the non-quasi-static (NQS) model and the parameters extracted from the 3D simulations (symbols) demonstrate excellent agreement. Thus, the proposed model is highly reliable for RF/millimeter-wave applications and hence suitable for providing information to RF engineers. All the parameters extracted in the linear ($V_d = 0.5\,V$) and saturation ($V_d = 1.0\,V$) regions are summarized in Table 2. Also, the effect of the metal workfunction engineering on all the parameters is summarized in Table 3.

4 Conclusions

Both the extrinsic and intrinsic parameters of the GEWE-SiNW MOSFET were directly extracted based on quantum simulations and compared with those of the SiNW and

conventional MOSFET in the GHz range to analyze their use in RF applications. The simulation results demonstrate a noticeable reduction in parasitic resistance, inductance, and capacitance for the GEWE-SiNW MOSFET compared with its conventional counterparts. Also, the reduction in the intrinsic C_{sdx} implies that the DIBL effect is less pronounced in the proposed device at HF. Moreover, by tuning the workfunction difference, the effect of parasitics, which hinder device performance at RF, can be decreased. Also, the admittance and impedance parameters are improved with incorporation of the GEWE scheme, due to reduced SCEs and maximum volume inversion owing to the cylindrical gate, thus providing necessary information to RF engineers for microwave communications. The modeling results match well with the Z and Y parameters extracted from 3D device simulations.

Acknowledgements The authors would like to thank the Microelectronics Research Lab, Department of Engineering Physics, Delhi Technological University (formerly DCE). N.G. is grateful to the University Grant Commission (UGC) for providing the financial assistance necessary to carry out this research work.

References

1. Deen, M.J., Fjeldly, T.A.: CMOS RF Modeling. Characterization and Applications. World Scientific, Singapore (2002)
2. Wood, J., Lamey, D., Guyonnet, M., Chan, D., Bridges, D., Monsauret, N., Aaen, P.H.: An extrinsic component parameter extraction method for high power RF LDMOS transistors. In: IEEE MTT-S International Microwave Symposium Digest, pp. 607–610 (2008)
3. Ytterdal, T., Cheng, Y., Fjeldly, T.A.: Device Modeling for Analog and RF CMOS Circuit Design. Wiley, New York (2000)
4. Cheng, Y., Deen, M.J., Chen, C.-H.: MOSFET modeling for RF IC design. IEEE Trans. Electron Devices **52**(7), 1286–1303 (2005)
5. Woerlee, P.H., et al.: RF-CMOS performance trends. IEEE Trans. Electron Devices **48**(8), 1776–1782 (2001)
6. Alam, M.S., Armstrong, G.A.: Extraction of extrinsic series resistance in RF CMOS. In: NSTI-Nanotech, pp. 136–139 (2004)
7. Buss, D.: Device issue in the integration of analog/RF functions in deep sub-micron digital CMOS. In: IEDM Technology Digest, pp. 423–426 (1999)
8. Kumar, A., Gupta, N., Chaujar, R.: Power gain assessment of ITO based transparent gate recessed channel (TGRC) MOSFET for RF/wireless applications. Superlattices Microstruct. **91**, 290–301 (2016)
9. Kumar, A., Gupta, N., Chaujar, R.: TCAD RF performance investigation of transparent gate recessed channel MOSFET. Microelectron. J. **49**, 36–42 (2016)
10. Subramanian, V., Abdelkarim, M., Bertrand, P., Morin, D., Guido, G., Willy, S., Stefaan, D.: Identifying the bottlenecks to the RF performance of FinFETs. In: IEEE 23rd International Conference on VLSI Design, pp. 111–116 (2010)
11. Cui, Y., Zhong, Z., Wang, D., Wang, W.U., Lieber, C.M.: High performance silicon nanowire field effect transistors. Nano Lett. **3**(2), 149–152 (2003)
12. Chen, X., Tan, C.M.: Modeling and analysis of gate-all-around silicon nanowire FET. Microelectron. Reliab. **54**(6–7), 1103–1108 (2014)
13. Iwai, H., et al.: Si nanowire FET and its modeling. Sci. China Inf. Sci. **54**(5), 1004–1011 (2011)
14. Wang, R., Zhuge, J., Huang, R., Tian, Y., Xiao, H., Zhang, L., Li, C., Zhang, X., Wang, Y.: Analog/RF performance of Si nanowire MOSFETs and the impact of process variation. IEEE Trans. Electron Devices **54**(6), 1288–1294 (2007)
15. Cho, S., Kang, I.M., Kim, K.R.: Investigation of source-to-drain capacitance by DIBL effect of silicon nanowire MOSFETs. IEICE Electron. Express **7**(19), 1499–1503 (2010)
16. Cho, S., Kim, K.R., Park, B.G., Kang, I.M.: RF Performance and small-signal parameter extraction of junctionless silicon nanowire MOSFETs. IEEE Trans. Electron Devices **58**(5), 1388–1396 (2011)
17. Long, W., Ou, H., Kuo, J.M., Chin, K.K.: Dual material gate (DMG) field effect transistor. IEEE Trans. Electron Devices **46**(5), 865–870 (1999)
18. Gupta, N., Chaujar, R.: Implications of transport models on the analog performance of gate electrode workfunction engineered (GEWE) SiNW MOSFET. In: IEEE 2nd International Conference on Devices, Circuits and Systems, pp. 1–5 (2014)
19. Gupta, N., Kumar, A., Chaujar, R.: Oxide bound impact on hot-carrier degradation for gate electrode workfunction engineered (GEWE) silicon nanowire MOSFET. Microsyst. Technol. **22**(11), 2655–2664 (2015)
20. Gupta, N., Kumar, A., Chaujar, R.: Impact of device parameter variation on RF performance of gate electrode workfunction engineered (GEWE)-silicon nanowire (SiNW) MOSFET. J. Comput. Electron. **14**(3), 798–810 (2015)
21. Gupta, N., Kumar, A., Chaujar, R.: TCAD analysis of frequency dependent intrinsic and extrinsic parameters of GEWE-SiNW MOSFET. In: Nanoelectronics, Materials & Devices, pp. 185–188 (2015)
22. Hu, C.: Modern Semiconductor Device for Integrated Circuit, pp. 261–274. Prentice Hall, Englewood Cliffs (2009)
23. ATLAS User's Manual: SILVACO Int. Santa Clara, CA (2014)
24. Gupta, N., Chaujar, R.: Influence of gate metal engineering on small-signal and noise behaviour of silicon nanowire MOSFET for low-noise amplifiers. Appl. Phys. A **122**(8), 1–9 (2016)
25. Iannaccone, G., Curatola, G., Fiori, G.: Effective Bohm Quantum potential for device simulators based on drift-diffusion and energy transport. In: Simulation of semiconductor processes and devices (SISPAD), pp. 275–278 (2004)
26. Suk, S.D., et al.: High performance 5 nm radius twin silicon nanowire MOSFET (TSNWFET): fabrication on bulk Si wafer, characteristics, and reliability. In: IEEE IEDM Technology Digest, pp. 717–720 (2005)
27. Yang, B., Buddharaju, K.D., Teo, S.H.G., Singh, N., Lo, G.Q., Kwong, D.L.: Vertical silicon-nanowire formation and gate-all-around MOSFET. IEEE Electron Device Lett. **29**(7), 791–794 (2008)
28. Rustagi, S.C., et al.: CMOS Inverter based on gate-all-around silicon-nanowire MOSFETs fabricated using top-down approach. IEEE Electron Device Lett. **28**(11), 1021–1024 (2007)
29. Polishchuk, I., Ranade, P., King, T.J., Hu, C.: Dual work function metal gate CMOS technology using metal interdiffusion. IEEE Electron Device Lett. **22**(9), 444–446 (2001)
30. Liu, J., Wen, H.C., Lu, J.P., Kwong, D.L.: Dual-work-function metal gates by full silicidation of Poly-Si with Co-Ni bi-layers. IEEE Electron Device Lett. **26**(4), 228–230 (2005)
31. Na, K.Y., Kim, Y.S.: Silicon complementary metal-oxide semiconductor field-effect transistors with dual work function gate. Jpn. J. Appl. Phys. **45**(12), 9033–9036 (2006)
32. Shirak, O., Shtempluck, O., Kotchtakov, V., Bahir, G., Yaish, Y.E.: High performance horizontal gate-all-around silicon nanowire field-effect transistors. Nanotechnology **23**, 395202 (2012)

33. Moon, D., Choi, S.-J., Duarte, J.P., Choi, Y.-K.: Investigation of silicon nanowire gate-all-around junctionless transistors built on a bulk substrate. IEEE Trans. Electron Devices **60**, 1355–1360 (2013)
34. Lee, J.-H., Kim, B.-S., Choi, S.-H., Jang, Y., Hwangc, S.W., Whang, D.: A facile route to Si nanowire gate-all-around field effect transistors with a steep subthreshold slope. Nanoscale **5**, 8968–8972 (2013)
35. Iwai, H.: Roadmap for 22 nm and beyond. Microelectron. Eng. **86**(7), 1520–1528 (2009)
36. Raghavan, A., Srirattana, N., Laskar, J.: Modelling and Design Techniques for RF Power Amplifiers. Wiley, New York (2007)
37. Lovelace, D., Costa, J., Camilleri, N.: Extracting small- signal model parameters of silicon MOSFET transistors. In: Proceedings IEEE MTT-S, pp. 865–868 (1994)
38. Tsividis, Y.: Operation and Modeling of the MOS Transistor, 2nd edn. Oxford Universiy Press, New York (2010)
39. Malik, P., Gupta, R.S., Chaujar, R., Gupta, M.: AC analysis of nanoscale GME-TRC MOSFET for microwave and RF applications. Microelectron. Reliab. **52**(1), 151–158 (2012)
40. Razavi, B., Yan, R.H., Lee, K.F.: Impact of distributed gate resistance on the performance of MOS devices. IEEE Trans. Circuits Syst. I Fundam. Theory Appl. **41**(11), 750–754 (1994)
41. Je, M., Shin, H.: Accurate four-terminal RF MOSFET model accounting for the short-channel effect in the source to drain capacitance. In: Simulation of Semiconductor Processes and Devices (SISPAD), pp. 247–250 (2003)

Springer

KSME

Journal of Mechanical Science and Technology 31 (1) (2017) 17~27
www.springerlink.com/content/1738-494x(Print)/1976-3824(Online)
DOI 10.1007/s12206-016-1203-5

Study of corrosive fatigue and life enhancement of low pressure steam turbine blade using friction dampers †

Loveleen Kumar Bhagi[1,*], Vikas Rastogi[2] and Pardeep Gupta[1]

[1]*Department of Mechanical Engineering, Sant Longowal Institute of Engineering and Technology, Longowal 148106, Punjab, India*
[2]*Department of Mechanical, Production, Industrial and Automobile Engineering, Delhi Technological University, Delhi 10042, India*

(Manuscript Received January 10, 2016; Revised July 8, 2016; Accepted July 16, 2016)

Abstract

The paper investigates the causes of failure of low pressure steam turbine blade of the last stage and suggests few techniques to overcome the failure causes is presented in this paper. The blade under investigation is made of chrome alloy steel. The fracture occurred at the airfoil region of the blade. The investigation included visual inspection, micro structural characterization, SEM-EDS microanalysis and spectroscopy test to identify the causes of failure. The paper also suggests the methods to reduce the blade fatigue subsequently through computation and to enhance the fatigue strength of the steam turbine blade. In order to reduce the blade fatigue susceptibility and to enhance the fatigue life of the steam turbine blade, the wedge shape friction damper is proposed.

Keywords: Steam turbine blade; Fracture; Corrosion; Fatigue; Friction damper; Bondgraph modeling

1. Introduction

Steam turbine is the simplest and most effective engine for converting large amounts of heat energy into mechanical work. During the past few decades, the operating temperatures of steam turbine engines have been on the rise to achieve higher and higher engine power and higher efficiency. This has necessitated a continuing advancement in high temperature withstand materials. The steam turbine blades are the critical components in main power equipment in nuclear and fossil power plants and in event of their failure, the power plant will shut down and cause long time power failure and economic loss.

As the high pressure steam from the boiler enters the turbine, it passes through different stages such as High pressure (HP), Intermediate pressure (IP) and Low pressure (LP) stage and with the help of turbine blades the linear motion of superheated steam gets converted into a rotary motion of the turbine shaft, which further drives the generator rotor [1]. As the steam expands, it acquires huge velocity and exerts a great force on the turbine blades [2]. In the course of steam expansion through LP turbine section, the steam generally crosses saturation line. This is the region where condensation of superheated steam begins and may be termed the Phase transition zone (PTZ) or Wilson line [3]. In steam turbines this phase transition zone is usually found at L-1 stage, which lies in the transonic flow region. Due to this transonic velocity (Mach = 1) the sonic shock waves is produced which, can be source of blade excitation and cyclic stresses that causing fatigue or corrosion fatigue [4]. Unfortunately, this is the region where maximum corrosion damage has been observed [5]. The general view of the double flow low pressure steam turbine blades is shown in Fig. 1, where L-1 stage blades are mentioned separately.

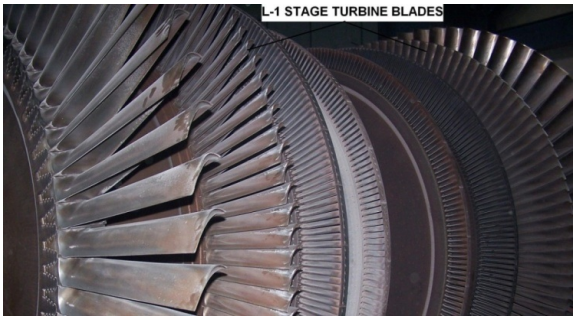


Fig. 1. L-1 row steam turbine blades of LP turbine stage.

The steam turbine blade failures can happen for many reasons including corrosion fatigue, fatigue fracture, stress corrosion, erosion, defective manufacturing and assembly, fault operation, design errors and others but it has been observed that in steam turbines the forced outages due to corrosion fatigue accounts for the most [6]. It is well-documented that LP steam turbine blades are more susceptible to failure than HP

*Corresponding author. Tel.: +91 9872650490
E-mail address: bhagiloveleen14@rediffmail.com
† Recommended by Editor Yeon June Kang
© KSME & Springer 2017



Fig. 2. Failed LP steam turbine blade from L-1 row.

and IP blades [7].

In this paper, the authors shall attempt to explore the root cause of failure of a last stage steam turbine blade of a 110 MW geothermal power plant. The premature failure of the blades actually occurred after a service life of 7 years and caused extensive damage to the power plant unit. The objective of this investigation was to identify the root cause of the failure and also suggest few corrective measures along with computational modeling, which would certainly be helpful in the fatigue life enhancement of the turbine blades.

2. Experimental methodology

The turbine blade under investigation is the failed steam turbine blade from L-1 row of LP section and was retrieved from the accident site (Fig. 2). From the manufacture's data, which is available in the power plant shows that the blade under consideration was made from chrome alloy steel X20Cr13 (AISI 420). The LP section has double flow composed of 8 stages. In May 1998, after completing approximately seven years of service, the unit was forced to shut down due to failure. While opening the turbine casing, it has been found that the blade from 24th stage or 3 A stage of LP rotor was fractured.

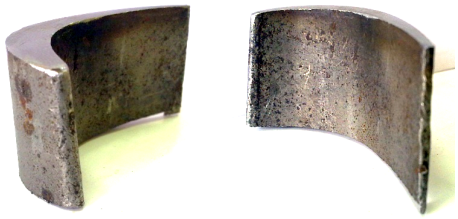
By following visual examination and digital photo documentation, portions of the fracture surface were cut for micro fractographic studies through scanning electron microscope (JEOL JSM-6610LV), which is equipped with an Energy dispersive spectroscopy (EDS) analysis facility. The specimens were subjected to SEM examinations as received condition from site after ultrasonic cleaning. All specimens were prepared for material characterization from the locations beneath the fracture surface. Chemical analysis was carried out using optical emission spectroscopy method. Specimen preparation for macroscopic and microscopic metallurgical evaluations on longitudinal and transverse sections was carried out using standard metallographic techniques, which was followed by etching with gycergia.

2.1 Visual inspection

The general orientation of the fracture surface is nearly normal to the longitudinal axis of the blade and it has been seen that the corrosion and pits were presented at the leading edge and trailing edge. The two sides of the failed blade, which bear different pressure levels i.e. Low pressure side and high pressure side, are visible in Figs. 3(a) and (b). Some portion of the blade was having dark brownish colour along with the pits, which clearly indicates the formation of oxides on the blade surface.

Table 1. Composition of the 13 % Cr turbine blade steel (Values are quoted as wt %).

	C	Si	Mn	Cr	S	P
Fractured blade	0.17	0.19	0.38	12.34	0.017	0.015
Standard AISI 420	0.16-0.25	Max 1.0	Max 1.5	12-14 Max	0.03 Max	0.04



(a) (b)

Fig. 3. Showing corroded turbine blade: (a) Low pressure side; (b) high pressure side.

2.2 Material characterization

For characterizing the material, first of all chemical analysis was conducted using optical emission spectroscopy (Optical emission analyzer spectrometer model-DV4) at the region beneath the fracture surface of the blade. The results of chemical analysis is given in Table 1, which shows the close agreement with the standard chemical composition for chrome alloy steel X20Cr13 (AISI 420).

2.3 Microstructure

For visualizing the microstructure of the fractured surface, the specimen was cut from the fractured blade airfoil region which was subsequently polished to take their optical micrographs. The micrograph clearly depicts a uniformly distributed hardened and tempered martensitic microstructure (Fig. 4), which shows that no micro-structural degradation has occurred at the fractured region of the turbine blade.

2.4 Fractography

The fractography of the fractured surface of blade airfoil region was examined using Scanning electron microscope (SEM), which was also equipped with EDS (JEOL JSM-6610LV). For analyzing the fractography, the failed blade was cut along its fractured region and the fractographs were taken at three different regions I, II and III which, is shown in Fig. 5. From the SEM fractograph of highlighted region 'I', it can be clearly observed that the crack has been initiated from this region. The beach marks presented in the fractograph shows that the crack propagates from the leading edge towards trail-

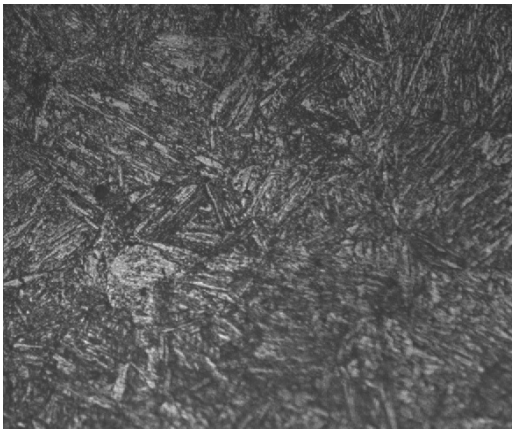


Fig. 4. Tempered martensitic microstructure of fractured blade airfoil regions (100 X).

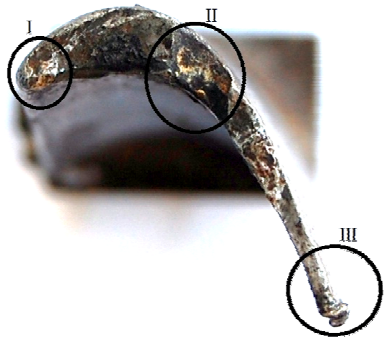


Fig. 5. General view of the fractured surface.

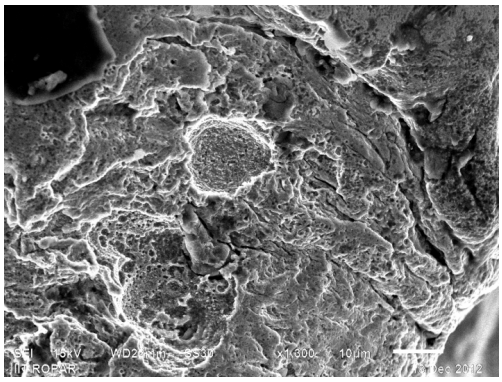


Fig. 6. SEM micrograph of region 'I' showing initiation and propagation of crack.

ing edge, which ultimately resulted into fatigue fracture (Fig. 6). Brittle fracture of transgranular type has also been observed in the fractographic view of region 'II' (Fig. 7). The beach marks along with the characteristic fatigue striations in the SEM fractograph of highlighted region 'III' can be seen in Fig. 8.

The presence of different nonmetallic elements in the fractured region as well as in the corrosion pits have been detected using EDS, equipped with SEM is shown in Fig. 9. The EDS spectrum thus obtained shows the presence of high percentage of Silicon along with other no-metallic elements such as oxy-

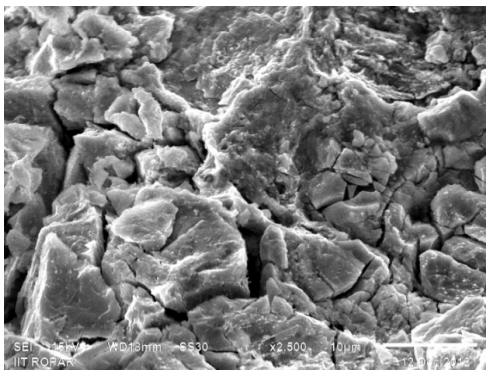


Fig. 7. SEM micrograph showing transgranular fracture along with secondary grain boundary cracks in the area of region 'II'.

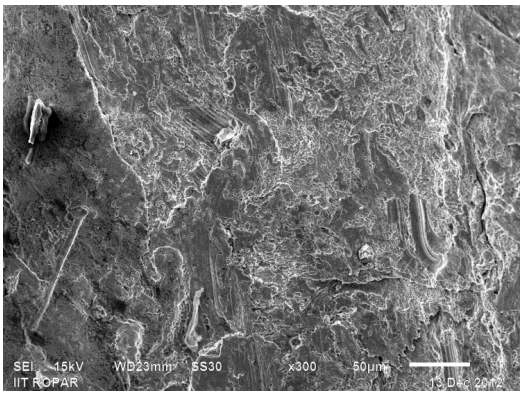


Fig. 8. SEM micrographs showing characteristic fatigue striations of region 'III'.

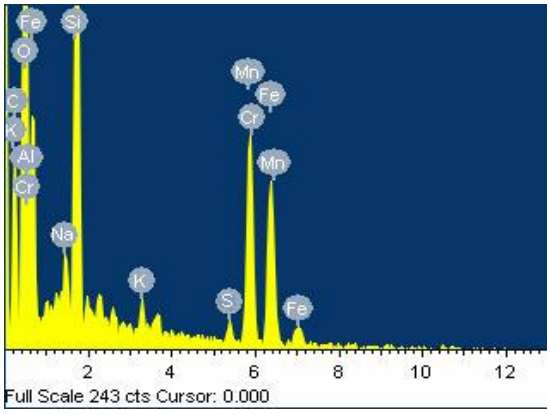
gen, potassium, aluminum, sodium, chlorine and sulphur.

3. Discussions and corrective measures

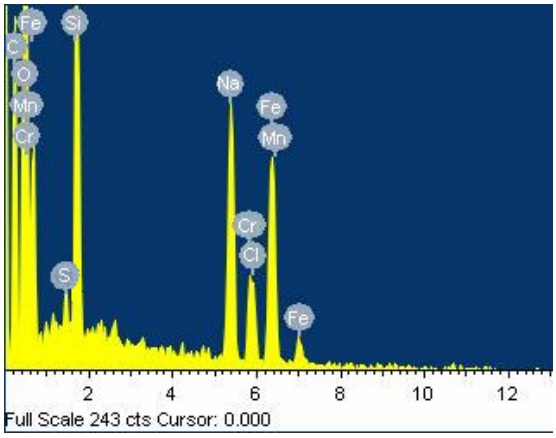
It has been found that the chemical composition and microstructure of fractured blade is in accordance with the standard martensitic stainless steel grade AISI 420. So, one may conclude that the blade under consideration was not failed due to any material defects.

From the SEM-EDS spectrum, it is observed that nonmetallic impurities are deposited on the fractured surface of the blade and corrosion pits. The major portion of these impurities contains Si and O, which may be due to the long term exposure in the atmosphere or may be due to the blade reaction with steam. The silicon oxide and chloride ion present in the steam, when strikes with the rotating turbine blades lead to the formation of corrosion pits on the blade surface and thus resulting to massive corrosion attack. These impurities generally foul the surface of stationary as well as moving blades and steadily build up during continuous running of a steam turbine in the Phase transition zone (PTZ). It may be due to the improper shut down and poor maintenance procedure, which contribute towards an aggressive environment that can lead to pitting and corrosion fatigue.

It can also be noticed from the SEM micrographs of the



(a)



(b)

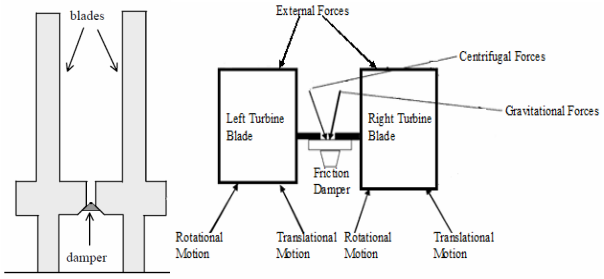
Fig. 9. EDS analysis results for the (a) surface of fractured blade; (b) corrosion pits.

failed blade surface, presented in Figs. 6-8 that well-defined striations or fatigue marks along with the trangranular fracture are observed on the fractured surface. It may be concluded from these fatigue marks and transgranular type nature of fracture that crack propagation was attributed due to fatigue resulting from the vibrations of the blade during startup and shut down operations of turbine.

All the above stated findings ultimately lead to the conclusion that the high corrosive environment along with high cycle fatigue stresses produced during the transient period of rotating blades had caused the low pressure steam turbine blade to break. Thus, corrosion fatigue is the main cause of forced outage in the turbine blade.

3.1 Corrective measures to counter corrosion fatigue

Low cyclic stresses and marginal steam chemistry acting together are the most frequent root causes in the last few stages of low pressure turbines. To avoid such type of failures in the steam turbine blades, the following corrective measures should be taken: (a) Get better steam chemistry [8]; (b) optimization or changing feed water and boiler water treatment to reduce concentration of impurities in steam [3]; (c) blade design in such way that friction damper at the blade root is considered [9, 10]; (d) eradicate the cause of blade excitation [11]; (e) minimize the mean stresses and alternating stresses by design [12].



(a) (b)

Fig. 10. (a) Blades with underplatform friction damper; (b) physical system of turbine blades with friction damper considering blade as a cantilever beam.

The fatigue failure can be avoided against large amplitude vibrations by the application of the dry fiction. The friction damper uses dry friction to dissipate energy and thus provides a passive, inexpensive and environmentally robust way to improve turbine blade life [13]. There are generally four types of damping used in the mechanical systems: (a) Viscous damping, (b) structural damping, (c) coulomb damping, and (d) friction damping (interfacial damping). General thumb rule for friction damping is "Larger the energy dissipation, larger is the effective damping in the system". To design friction dampers we have generally depends upon experimental data or precedent experience, since friction behaves non-linear in nature so its mathematical formulation and determination of dynamic response become complicated. To solve this problem we use bond graph technique to model and simulate the turbine blade with friction damper system.

4. Methodology of friction damper to enhance fatigue life

Blade vibrations are generally excited by transient fluid flow conditions and may become dangerous during starting and shut downs operations of the turbine. Under these transient conditions of operation the high dynamic stresses generated which, lead to high cycle fatigue problems in the blade. The source and magnitude of the excitation and the damping in the blade [14] have shown significant effect on the dynamic stresses. In the case of a free-standing turbine blade, damping has been shown to be inherently low [15] which, can result in high dynamic stresses at resonance conditions. Freestanding turbine blades have characteristically low structural damping and thus only way to reduce these high dynamic stresses can done by the use of additional friction damping devices, such as underplatform friction dampers.

The physical system of the turbine blade with wedge shaped underplatform friction damper has been shown in Figs. 10(a) and (b). The under platform friction dampers are metal parts,

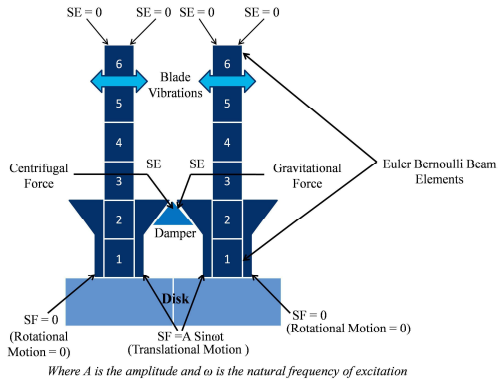


Fig. 11. Physical system of the turbine blade with friction damper.

which are placed into a cavity between two contiguous blades underneath the blade airfoils and pressed against the blades by the action of centrifugal force [16]. Because of high rpm of turbine rotor the centrifugal force acting on the damper mass is large enough to keep the damper in contact with the platform surfaces at all times. Blade vibrations under different working conditions can be analyzed numerically by simulating the model of the fractured blade. The role of friction dampers is very significant in vibrating structures in order to increase damping and reduce vibratory stresses by dissipating energy.

5. Dynamic modeling of underplatform friction damper for life enhancement of turbine blade

5.1 Physical system

For the present work, steam turbine blade has been taken as an Euler-Bernoulli beam in the form of six reticules and friction damper is placed in between the recess of two blades (Fig. 11). In this model, the friction damper is used after 2nd reticule of the E-B beams. Bondgraphs may be used efficiently and effectively to model the system and getting the simulation results for analysis of fatigue life enhancement of steam turbine blade as a basis for comparison.

5.2 Bondgraph modeling of turbine blade considering Euler-Bernoulli beam model

For the modeling of Euler-Bernoulli beam, one end of beam is fixed (Fig. 12) the effect of rotary inertia is neglected compared with those of the linear inertia (m, the mass per unit length). The deformations related with transverse shear are not included as well. These assumptions are reasonable provided that the beam is thin and, for vibratory behavior, that the effective wave lengths of the individual modes are sufficiently large compared with the thickness.

We assume, the mass density of Euler-Bernoulli beam is ρ and cross sectional area is A are constant along the beam, the Young's modulus, E, and the area moment of inertia, I. The product of EI gives the flexural rigidity or bending stiffness of

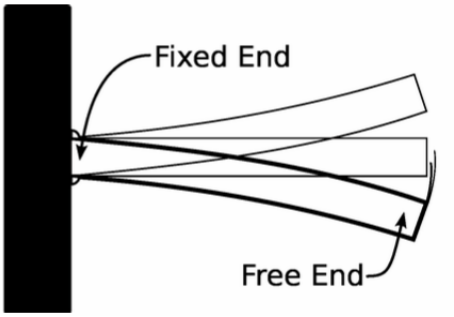


Fig. 12. A beam with one end is fixed.

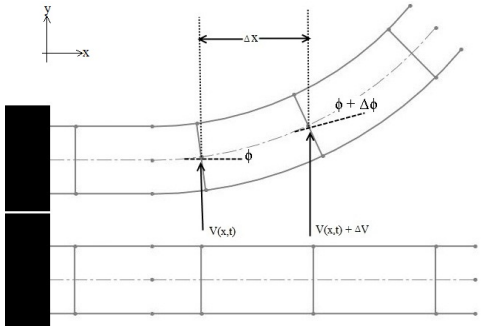


Fig. 13. Euler-Bernoulli beam.

the beam. Let us consider small element of beam of having length Δx. Suppose there is an upward force, w(x)/per unit length, acting on the beam which produces transverse displacement V(x, t) and V (x, t) + ΔV at position x and Δx is shown in Fig. 13. Q(x) and M(x) are shear force and bending moment acting on the finite element of the beam.

Since the modeling of distributed parameters systems typically begin with finite approximation and hence space reticules, one may proceed with the governing equations in difference form. To derive the equations of motion for the beam let us consider a generic differential element of length Δx.

The forces acting on an infinitesimally small lump of length Δx in y-direction is:

$$\sum F_y = 0$$
$$\sum M_y = 0 .$$

For shear force

$$-Q(x) + Q(x + \Delta x) = -w(x)\Delta x$$
$$\underset{\Delta x \to 0}{lt} \frac{Q(x + \Delta x) - Q(x)}{\Delta x} = -w(x)$$
$$\frac{dQ}{dx} \cong w(x) . \quad (1)$$

For bending moment

$$-M(x) + Q(x)\frac{\Delta x}{2} + M(x + \Delta x) + Q(x + \Delta x)\frac{\Delta x}{2} = 0$$

$$\lim_{\Delta x \to 0} \frac{M(x+\Delta x) - M(x)}{\frac{\Delta x}{2}} = -2Q(x)$$

$$Q(x) = -\frac{dM}{dx}. \tag{2}$$

From Eqs. (1) and (2) we get

$$w(x) = -\frac{d^2M}{dx^2}. \tag{3}$$

From Eq. (3) we can see that for constant load, w(x) the shear force, Q(x), varies linearly and then M(x) will be in a quadratic form.

Kinematic equation:

The angle ϕ is the angular deformation due to bending about z-axis.

$$\tan\Phi \approx \Phi = \frac{\Delta V(x,t)}{\Delta x} << 1\ ;\ \frac{d\Phi}{dx} = \frac{d^2V(x,t)}{dx^2}.$$

Constitutive equation:

Bending moment,

$$M = \int_A -y\sigma dA. \tag{4}$$

Modified form of Eq. (4) will be

$$M = E\frac{d^2V(x,t)}{dx^2}\int_A y^2 dA$$

$$M = EI\frac{d^2V(x,t)}{dx^2}. \tag{5}$$

After putting the value of M from Eq. (5) in Eq. (3), we get

$$w(x) = EI\frac{d^4V(x,t)}{dx^4}. \tag{6}$$

For the present theories, the statement of Newton's Second Law in the transverse direction may be written as

$$w(x,t)dx + \left[Q(x,t) + \frac{\partial Q}{\partial x}dx\right] - Q(x,t) = m(x)dx\frac{\partial^2 V(x,t)}{\partial t^2}$$

which reduces to

$$w(x,t) + \frac{\partial Q(x,t)}{\partial x} = m(x)\frac{\partial^2 V(x,t)}{\partial t^2}$$

$$\rho A\frac{\partial^2 V(x,t)}{\partial t^2} + EI\frac{\partial^4 V(x,t)}{\partial x^4} = w(x,t). \tag{7}$$

The above relation which is shown by the Eq. (7) is the transverse motion of Euler-Bernoulli beam in the form of differential equation.

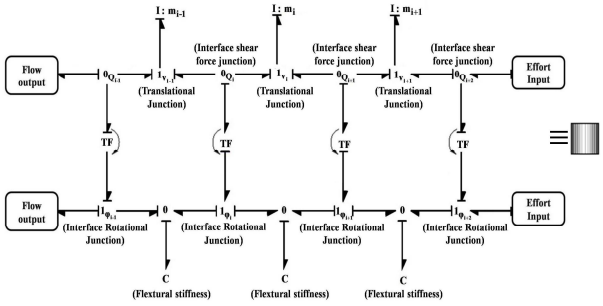


Fig. 14. Bond graph finite-lump model of reticulated Euler-Bernoulli beam.

We can use above differential equation of motion for displacements of each reticule of an overall beam for the construction of the mass matrix and the stiffness matrix in the dynamic equation of the entire beam in a systematic manner from the corresponding matrices of the reticules.

Before drawing the bond graph model we have first divided the E-B beam into the reticules. The shear forces are acting at the boundaries of each reticule and represented by 0-junctions. The bond graph model of E-B beam has been created from the above kinematic and constitutive equations (Fig. 14).

The Euler Bernoulli beam as turbine blade is modeled with six reticules (Beam elements). We can add more number of reticules to improve the modeling accuracy but this will lead to increase the computational amount significantly.

6. Simulation study

The bond graph model of the turbine blade with friction damper is shown in Fig. 15 and simulated using a Runge-Kutta fourth order method on the software SYMBOLS Sonata® [17]. The shaft is assumed to be rotated at constant speed. The bond graph model consists of solid rotating discs with blades, which is discretized into six reticules with one end attached to the discs with one end free similar to cantilever beam. The parameters used for simulation are presented in Table 2. The dimensional data are as follows: Lbeam = 1.2 m, material and blade properties are as follows: E = 1.93e11 N/mm^2, ρ = 7850 kg/m^3. All the parameters presented in the Table 2 are used for simulation. The results of simulation are presented in the next sub-section.

Referring Fig. 16, it can be seen that damper mass has significant effect of on the frictional force between blade and damper. To design optimum damper configuration, the value of friction forces should not large enough, it will stick, restraining the relative motion between the blade and damper and if the value of friction force is too small then it will not dissipates sufficient amount of vibratory energy from the system. From this figure it is clear that between these two extreme values of friction forces the damper mass is optimal.

Energy dissipation due to high contact pressure and micro slipping friction in blade root by means of underplatform friction damper is useful to introduce damping in the system.

Table 2. Parameters and values used for simulation.

	Parameter	Value
For turbine blade material	Length of each capsule	0.2 m
	Modulus of elasticity	1.93 e11 N/m^2
	Density	7850 kg/m^3
	Radius of the turbine blade	0.4 m
	Moment of inertia	1 e-7 m^4
	Internal damping co-efficient	2.50x10^{-4} Ns/m
For friction damper	Surface area of the damper	0.003 m^2
	Damping coefficient of the friction	0.7 N.m.s
	Mass of friction damper	0-0.9 Kg
	Inertia of the damper	4.17 e-4 m^4
	Frequency of turbine blade	200 Hz
	Elasticity of the material of the damper	1.78 e11 N/m^2
	Second moment of area of damper	1.25 e-8 m^4
	Frequency of excitation	50 Hz
	Amplitude of excitation	0.2 m
	Reynolds number	10000
	Inertia distance of contact	0.01 m

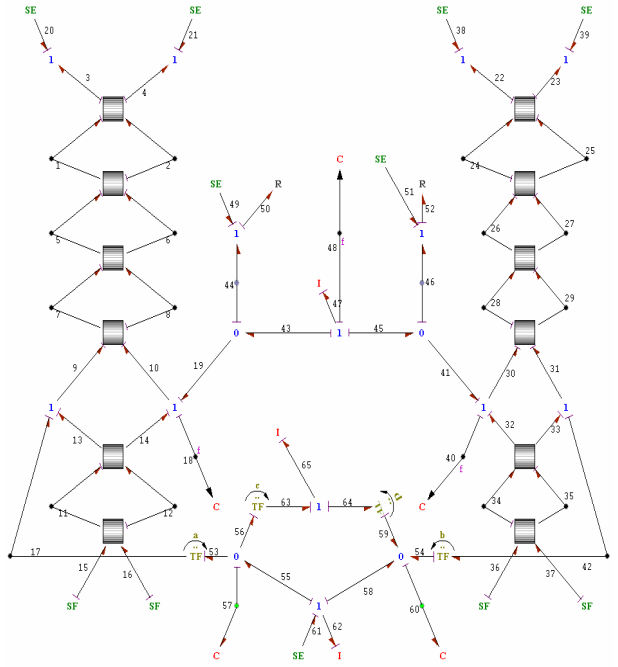


Fig. 15. Integrated bondgraph model of turbine blade with friction damper.

During the relative motion between the contacting surface of damper and platform, energy is dissipated due to friction. The results were obtained for energy dissipation between turbine blade and damper for varying damper mass is shown in Figs. 17(a)-(c) and it is concluded that maximum energy dissipates from the system occurs for 0.7 kg of damper mass.

Energy dissipation due to friction induced is further depending upon the normal contact force between the turbine blade

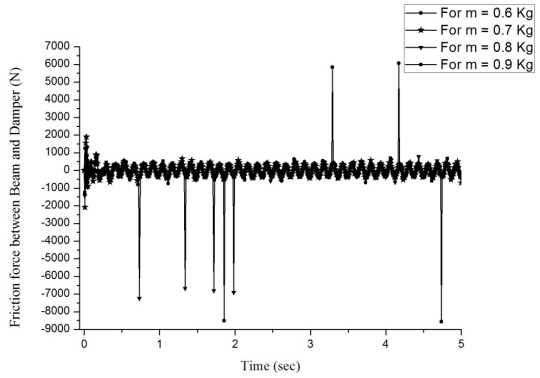
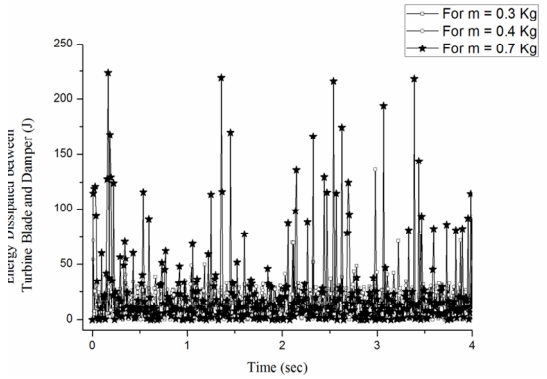
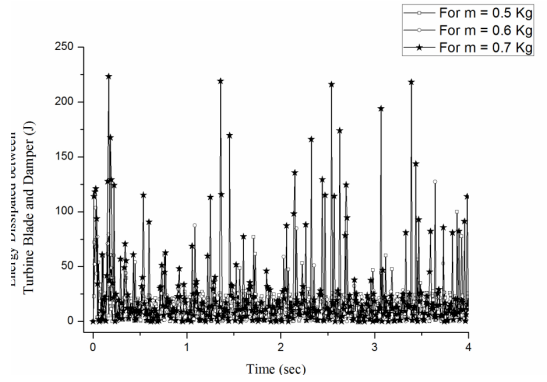


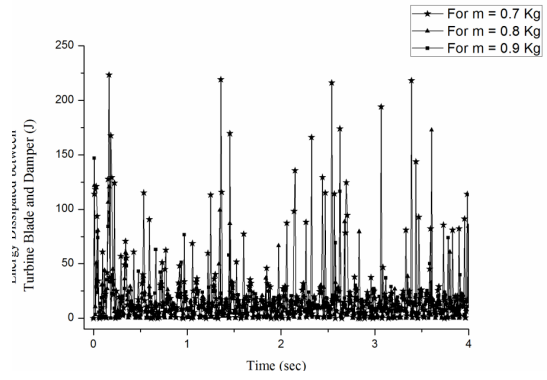
Fig. 16. Frictional force between beam and damper.



(a)

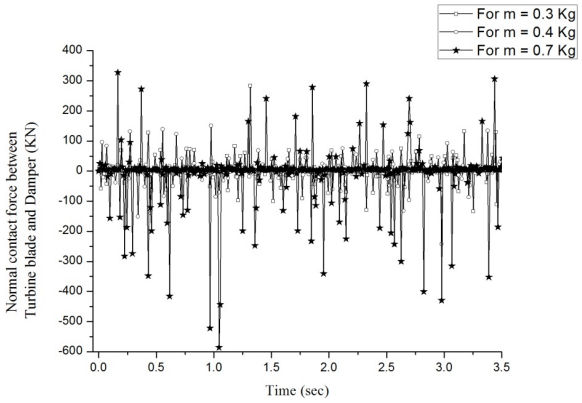


(b)

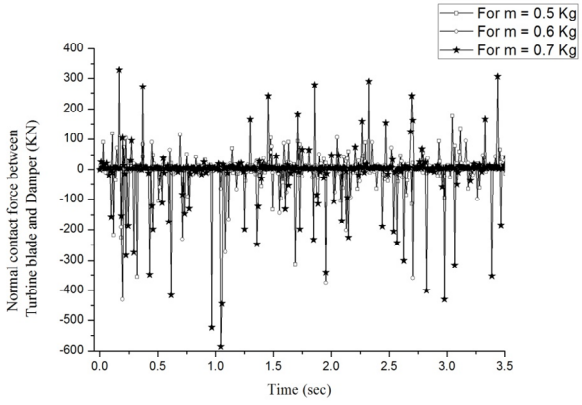


(c)

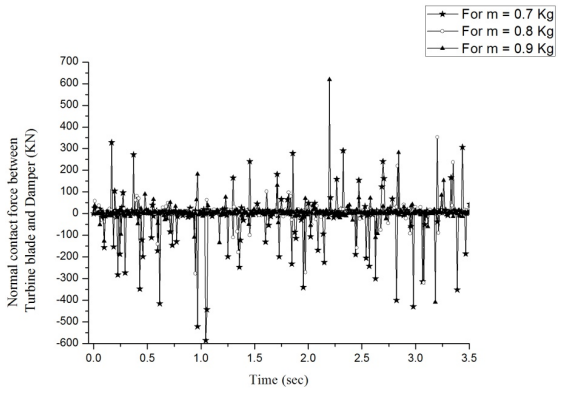
Fig. 17. Energy dissipated between blade and damper.



(a)



(b)

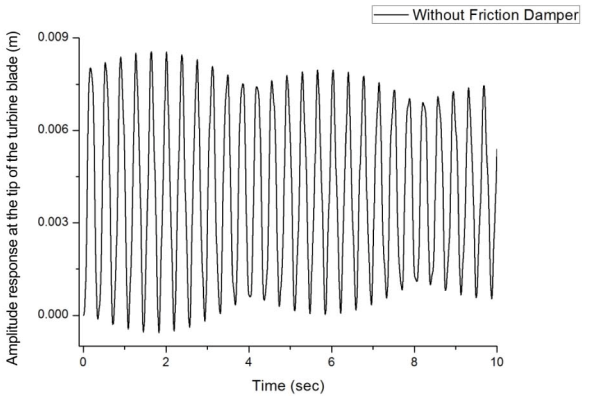


(c)

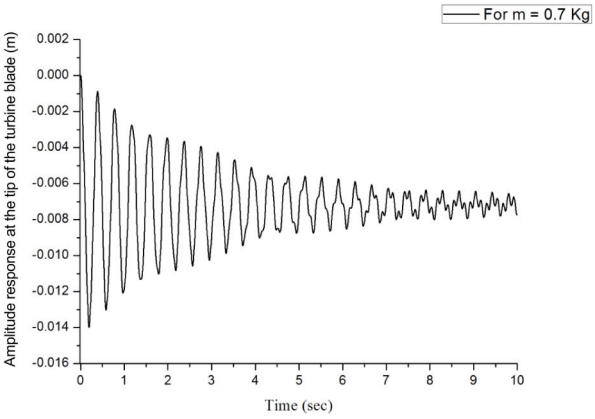
Fig. 18. Normal contact force between Turbine blade with varying damper mass.

and damper. From the Figs. 18(a)-(c), it is observed that normal contact force varies with the change in the mass of damper and at 0.7 kg of damper mass the normal contact force varies considerably.

From the Fig. 19(a), without introduce the friction damper between the blades we can see, the response is not dying out and this will leads to accumulate high dynamic stresses in the low pressure steam turbine blades which, may ultimately result into the catastrophic failure in the steam turbine. With the use of optimum friction damper configuration it is clearly seen



(a)



(b)

Fig. 19. Displacement of beam reticule (a) without friction damper; (b) with damper mass 0.7 kg.

(Fig. 19(b)) that there is an exponential decay in the transient response. Friction damper facilitates in reducing the vibratory stresses at turbine blades and hence protected the turbine blades from high cycle fatigue stresses. The optimum damper mass for the L-1 low pressure steam turbine blade model recommended by the author in this paper is 0.7 kg and this will further leads to the fatigue life enhancement of the steam turbine blades.

Simulation results showed a great agreement to the other parameters/values found. This unique model of dry friction damper as suggested may be easily applied in power plants to increase the productivity and to reduce failures of blades.

7. Conclusions

The purpose of this paper was to generate a Bond graph (BG) model for a steam turbine blade system. This is done by first making a BG model for Turbine blade by considering blade as an Euler-Bernoulli beam, which, is then incorporated in the integrated BG model of steam turbine blade with friction damper. The result obtained from the modeling and simulation of the integrated BG model is more convincing with the results found by other authors working in the field of turbine blade failure.

Based upon the all fractographic and computational results, the following conclusions have been made:

- According to the microstructure and hardness analysis it is concluded that the steam turbine blade under investigation was not failed due to its material defect.
- The oxides of sodium and silicate were detected on the fractured surface, which leads to the formation of corrosion pits and responsible for massive corrosion attack.
- The transgranular fracture mode and the fatigue marks (beach marks) show that the cause of failure is corrosion fatigue, which has reduced the overall working life-span of turbine blade.
- The turbine blade may be designed with effective friction damping and the BG computational modeling has been incorporated to show the efficacy of the friction damper for life enhancement techniques.
- The high damping coefficient material may be used for the manufacturing of LP steam turbine blades so that the energy of vibration may be dissipated.
- The Optimum friction damper required to reduce the vibratory stresses and hence to increase the fatigue life of turbine blade, comes out to be 0.7 kg.

References

[1] E. Poursaeidi and M. R. M. Arhani, Failure investigation of an auxiliary steam turbine, *Eng. Fail Anal.*, 17 (2010) 1328-1336.

[2] I. I. Kirillov and R. M. Yablonik, *Fundamentals of the theory of turbines operating on wet steam,* Leningrad: Mashinostroyeniye Press (1968).

[3] O. Jonas and B. Dooley, Major turbine problems related to steam chemistry: R&D, Root causes, and Solutions, *Fifth international Conference on Cycle Chemistry in Fossil Plants*, EPRI, Palo Alto, California (1997).

[4] M. Stastny, O. Jones and M. Sejna, Numerical analysis of the flow with condensation in a turbine cascade, *Report on Moisture Nucleation in Steam Turbines*, EPRI, Palo Alto, California; TR-108942 (1997).

[5] S. L. Alexander, *Steam Turbines for Modern Fossil-Fuel Power Plants,* The Fairmont press, Lilburn, Georgia (2008).

[6] L. K. Bhagi, V. Rastogi and P. Gupta, Fractographic investigations of the failure of L-1 low pressure steam turbine blade, *Case Studies in Eng. Fail Anal.*, 1 (2013) 72-78.

[7] N. K. Mukhopadhyay, S. G. Chowdhury, G. Das, I. Chattoraj, S. K. Das and D. K. Bhattacharya, An investigation of the failure of low pressure steam turbine blades, *Eng. Fail Anal.*, 5 (1998) 181-193.

[8] G. Das, S. G. Chowdhury, A. K. Ray, S. K. Das and D. K. Bhattacharya, Turbine blade failure in a thermal power plant, *Eng. Fail Anal.*, 10 (2003) 85-91.

[9] V. Rastogi, V. Kumar and L. K. Bhagi, Dynamic modeling of underplateform damper used in turbomachinery, *J. World Academy of Sc. Eng. and Tech.*, 61 (2012) 536-545.

[10] J. H. Griffin, A review of friction damping of turbine blade vibration, *International Journal of Turbo and Jet Engines*, 7 (1990) 297-307.

[11] M. P. Singh, J. J. Vargo, D. M. Schiffer and J. D. Dello, SAFE diagram-A design and reliability tool for turbine blading, *Seventeenth Turbomachinery Symposium*, Dallas, Texas (1988).

[12] J. S. Rao and A. Saldanha, Turbomachine blade damping, *J. of Sound and Vib.*, 262 (2003) 731-738.

[13] L. K. Bhagi, V. Rastogi, and P. Gupta, A brief review on modeling approaches of friction dampers used in turbomachinery, *International Conference on Soft Computing for Problem Solving*, IIT Roorkee, Roorkee, India (2011).

[14] J. S. Rao and N. S. Vyas, Determination of blade stresses under constant speed and transient conditions with nonlinear damping, *J of Eng. for Gas Turbines and Power*, 118 (1996) 424-433.

[15] N. F. Rieger, *Damping properties of steam turbine blades, Rotordynamics 2: Problems in turbomachinery*, Edited by N.F. Rieger, International Centre for Mechanical Sciences, Italy (1988).

[16] C. M. Firrone and S. J. Zucca, Underplatform dampers for turbine blades: The effect of damper static balance on the blade dynamics, *Mechanics Research Communications*, 36 (2009) 515-522.

[17] A. Mukherjee and A. K. Samantaray, *SYMBOLS-shakti® User's Manual,* High-tech consultants, STEP, Indian Institute of Technology, Kharagpur, India (2005).

Appendix

A. An introduction to bond graph modelling

A.1 Basic concept

Bond graph is an explicit graphical tool for capturing the common energy structure of systems. It increases one's insight into systems behaviour. In the vector form, they give concise description of complex systems. Moreover, the notations of causality provide a tool for formulation of system equations. Bond graphs are a domain-independent graphical description of dynamic behaviour of physical systems. This means that systems from different domains (electrical, mechanical, hydraulic, acoustical, thermodynamic, and material) are described in the same way. Bond-graph modelling is a powerful tool for modelling engineering systems, especially when different physical domains are involved.

A.2 Power variables of bond graphs

As energy can flow back and forth between two power ports of different nodes, a half arrow is added to each bond indicating a reference direction of the energy flow.

Furthermore, it is a useful modelling construct to represent the amount of power, P(t), at each time (t) by the product of two conjugate variables, which are called effort (e) and flow (f) respectively.

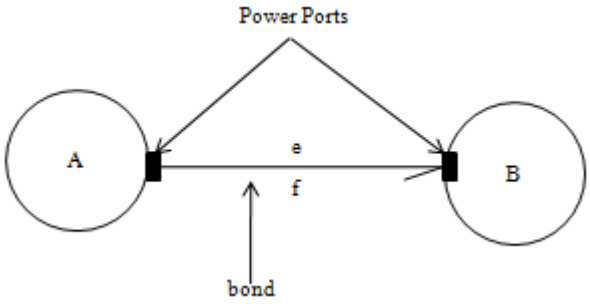


Fig. A.1. The energy flow between two submodels represented by a bond.

$P(t) = e(t).f(t)$.

A.2.1 Ports and bonds

The contact point of a submodel where an ideal connection will be connected is called a power port or port for short. Th connection between two submodels is called a power bond or bond; it is drawn as a single line (Fig. A.1). This bond denotes an ideal energy flow between the two connectedsubmodels. The energy entering the bond on one side immediately leaves the bond at the other side (power continuity).

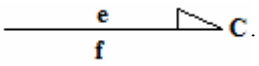
The energy flow along a bond has the physical dimension of power, being the product of two variables. In each physical domain, there is such a combination of variables, for which a physical interpretation is useful.

A.3 Standard elements

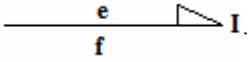
A.3.1 Bond graph elements (one port elements)

Bond graph elements are drawn as letter combinations (mnemonic codes) indicating the type of element. The bond graph elements are:

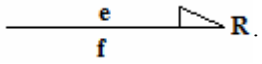
C - Capacitive or storage element for state variable of the system, e.g. capacitor (stores charge), spring (stores displacement). The bond graphic symbol for a C-element is shown below. This device stores and gives up energy without loss.



I- Inertial or storage element for a p-type state variable e.g. inductor (stores flux linkage), mass (stores momentum). The inertial element is used to model inductance effects in electrical systems and mass or inertia effects in mechanical or fluid systems. The bond graph symbol for an inertial element is depicted in the figure given below.



R- is a resistive element. Which dissipates free energy e.g., electric resistor, mechanical damper. The bond graph symbol for the resistive element is shown below.

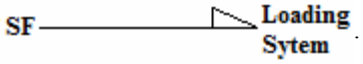


SE and SF - Effort and flow Sources like electric mains (voltage source), gravity (force source), and pump (flow source) and other analogous active elements.

The effort source is represented by

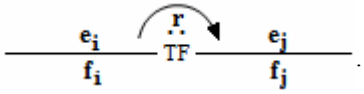


and the flow source is represented by

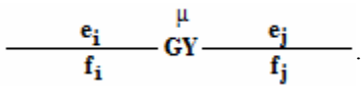


A.3.2 Bond graph elements (two port elements)

TF- Transformer, e.g. an electric transformer, toothed wheels, lever. An ideal transformer is represented by TF and no power is stored or dissipated in this device. Efforts are transduced to efforts and flows to flows. The parameter 'r' is the proportionality constant and is called the transformer module or ratio. The bondgraph element of transformer is shown below.



GY - Gyrator, e.g. electromotor, centrifugal pump. Unlike transformer, a gyrator establishes relationship between flow to effort and effort to flow. The simplest gyrator is a mechanical gyroscope. The parameter μ is the gyrator ratio.





Loveleen Kumar Bhagi has done his B. Tech. (2000) from Punjab Technical University, Jalandhar, Punjab, India and he received his Master degree with distinction on Manufacturing Engineering from Sant Longowal Institute of Engineering and Technology, Longowal, Sangrur, Punjab, India, in 2009. He is a member of ISTE, SAE and IAENG. Presently he is a Professor (Associate) and Head of Mechanical Engineering Department at Aryabhatta Group of Institutes, Barnala, Punjab, India. His main research interests include Turbomachinery blade vibrations and corrosion fatigue failure.



Vikas Rastogi obtained his Master of Technology degree from National Institute of Technology, Kurukshetra in 1995 and Ph.D. from Indian Institute of Technology, Kharagpur. He is currently a Professor and a Ph.D. candidate supervisor at Mechanical, Production, Industrial and Automobile Engineering Department, Delhi Technological University, India. His main research interests include Lagrangian-Hamiltonian mechanics for general class of systems, study of Symmetries for discrete and continuous systems, Bond Graph modeling and Simulation, rotor dynamics.



Pardeep Gupta obtained his Master of Technology degree from Punjab Engineering College, Chandigarh in 1997 and Ph.D. from National Institute of Technology, Kurukshetra. He is currently a Professor and Head at Mechanical Engineering department, Sant Longowal Institute of Engineering and Technology, Longowal, Sangrur, Punjab, India. His main research interests include production and industrial engineering.

Optical Materials 64 (2017) 250–256

Contents lists available at ScienceDirect

Optical Materials

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

Synthesis and optical properties of core-multi-shell CdSe/CdS/ZnS quantum dots: Surface modifications

R.K. Ratnesh, Mohan Singh Mehata*

Laser-Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

ARTICLE INFO

Article history:
Received 27 October 2016
Received in revised form
23 November 2016
Accepted 26 November 2016

Keywords:
Core-multi-shell Q-dots
CdSe/CdS/ZnS
Synthesis
Optical and structural properties
Lifetime

ABSTRACT

We report two port synthesis of CdSe/CdS/ZnS core-multi-shell quantum dots (Q-dots) and their structural properties. The multi-shell structures of Q-dots were developed by using successive ionic layer adsorption and reaction (SILAR) technique. The obtained Q-dots show high crystallinity with the step-wise adjustment of lattice parameters in the radial direction. The size of the core and core-shell Q-dots estimated by transmission electron microscopy images and absorption spectra is about 3.4 and 5.3 nm, respectively. The water soluble Q-dots (scheme-1) were prepared by using ligand exchange method, and the effect of pH was discussed regarding the variation of quantum yield (QY). The decrease of a lifetime of core-multi-shell Q-dots with respect to core CdSe indicates that the shell growth may be tuned by the lifetimes. Thus, the study clearly demonstrates that the core-shell approach can be used to substantially improve the optical properties of Q-dots desired for various applications.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Quantum dots (Q-dots)/semiconductor nanocrystals (NCs) and its core-shell structure have great potential applications. For examples, used as diagnostic and imaging agents in biomedicine and as semiconductors in the electronics industry [1–4]. As compared to organic dyes, Q-dots exhibit attractive property, a very broad absorption band and strong photoluminescence (PL). The CdSe and core-multi-shell structures Q-dots are widely used because of quantum confinement effect of electron and their emission wavelengths span in the entire visible spectrum [5–7]. One of the key features of colloidal semiconductor Q-dots is their tunability and narrow PL band, which depends on particle size, their distribution and surface chemistry. Though PL quantum yield (QY), decay kinetics and stability are closely related to the number of dangling bonds at the core particle surface. Therefore, surface modification by inorganic passivation layers and organic capping ligands is usually employed to optimize the required parameters [8].

The molecular surface modification of Q-dots is required to improve the PL efficiency [9,10], colloidal stability and also develop a reliable processing chemistry [11,12]. It is difficult simultaneously to passivate both anionic and cationic surface sites by organic ligands. Therefore, dangling bonds are some extent remain on the surface [13]. To completely passivate the NCs, various inorganic surface modification techniques have been developed [14,15]. Mainly, CdSe NCs were covered either with CdS or with ZnS [16,17]. To prepare a core/shell structure the band gap of the core should lies energetically within the band gap of the shell material, and the photo-generated electrons and holes are mainly confined inside the CdSe. These core-shell NCs are found to be more robust against chemical degradation or photo-oxidation [16].

The specific requirements for preparation of ideal core/shell NCs are to produce particles of high crystallinity the core and shell material should have similar lattice parameters such that the shell growth happens in an epitaxial manner [16] and the shell material should possess a much higher band gap than the core to suppress tunneling of the charge carriers from the core to the newly formed surface atom of the shell [18]. Several group have prepared CdSe/CdS, CdSe/ZnS and CdSe/CdS/ZnS core/shell Q-dots using cadmium oxide (CdO), zinc hydroxide ($Zn(OH)_2$), trioctylphosphine oxide (TOPO) and hexadecyl amine (HAD), and obtained the PL QY of about 30–50% [10,16–20]. Talpin et al. [21] have prepared various core/shell Q-dots using dimethyl cadmium acetate $Cd(CH_3COOH)_2$ and trioctylphosphine selenide (TOPSe) in which hydrogen sulfide (H_2S) used as sulfur (S) source and HDA–TOPO-TOP (trioctylphosphine) as stabilizing mixture [21]. Though the TOPO based Q-dots are considered as more toxic and create adverse effects on living organism.

* Corresponding author.
E-mail addresses: msmehata@gmail.com, mohan.phy@dce.edu (M.S. Mehata).

http://dx.doi.org/10.1016/j.optmat.2016.11.043
0925-3467/© 2016 Elsevier B.V. All rights reserved.

To use the core-shell Q-dots for biological applications, Q-dots should be water soluble because almost all biological environments are considered as water. The water soluble Q-dots can be utilized directly as labeling and imaging for biological applications [22], which can be synthesized by different methods. The hot-injection approach is used to prepare Q-dots using high-temperature inorganic solvents such as TOPO and HDA. These solvents are hydrophobic in nature, however, by replacing the hydrophobic TOPO and TOP layer with bifunctional molecules containing thiol such as mercaptopropionic acid (MPA) a water soluble Q-dots can be prepared. One end of the thiol (MPA) binds onto the surface of the Q-dots, and the carboxyl moiety confers towards water solubility. It has been reported [23–27] that the thiol-capped/functionalized Q-dots exhibits a higher stability and their terminal functionalities can bind with other molecules which, make the Q-dots suitable for various applications. The thiol-based ligand is used to modify the surfaces of Q-dots. However, thiols quench the PL of Q-dots [28–30]. The ligand-exchange reactions have become one of the most popular methods used for the surface modification of Q-dots, which involves the dynamic substitution of stabilizing ligands under certain conditions. This thiolates exchange approach has been used to change hydrophobic nature of magnetic and semiconducting nanoparticles into hydrophilic [31–35].

In this paper, we report the TOPO free synthesis of CdSe (core) and CdSe/CdS/ZnS (core-multi-shell) Q-dots. The core-shell structure is prepared by hot injection SILAR technique and examined using various techniques. The synthesized Q-dots are controllable in size/shape, crystallinity and grown in an epitaxial manner. The quantum yield (QY) of core-shell Q-dots is about 40% and for water soluble it various from 8 to 15% depending on the pH.

2. Experimental section

2.1. Chemicals

Trioctylphosphine (TOP, 90%), Oleic acid (OA, ≥90%), selenium powder (99.5%), zinc oxide (ZnO, 99.99%), cadmium oxide (CdO, >99%), sulfur (≥99), hexadecyl amine (HDA, 90%) and mercaptopropionic acid (MPA) were purchased from Aldrich Co. 1- Octadecene (ODE, ≥95.0%) was purchased from Alfa Aesar. Toluene (99.5%) and methanol (99.93%) were of spectroscopy grade and used without further purification.

2.2. Synthesis of CdSe core Q-dots

The synthesis was carried out under an inert atmosphere using nitrogen (N_2) gas. In order to provide rigorously oxygen and water free environment for the synthesis, all the substances were dried and degassed before used. Initially, the solutions of the Cd and Se precursors were prepared separately [22]. To make Se precursor, 0.0205 gm Se powder, 0.12032 ml TOP and 2 ml octadecane were mixed and put the mixer on the hot plate stir for around one hour. Similarly, for the Cd precursor, a mixer of 0.102 gm CdO, 1.23 ml oleic acid and 14 ml ODE was taken and then the mixture (CdO-OA-ODE) was transferred into three neck flask and put into heating rotamental and stirred continuously at 180 °C until the solution becomes colorless. Upon increasing temperature, up to 240 °C, 2 ml Se stock solution was transferred to the three-neck flask and kept for 60 s. After that, heating mantle was removed, the temperature of the reaction mixture (a mixture of Cd and Se precursor) was suddenly decreased to the room temperature, 20 ml of methanol was added to the reaction mixture and finally centrifuges the obtained solution at 6500 rpm for 10 min. The process is repeated 2–3 times to obtained pure CdSe Q-dots, and then the collected eluent was dispersed into toluene for further investigation.

2.3. Preparation of the precursor solution for the SILAR method

Shell growth around the core was performed using SILAR method [36]. For preparation of precursor solutions, metal oxide (ZnO and CdO), elemental sulfur and selenium were used. Zinc precursor solution (0.1 M) was prepared by dissolving ZnO (0.1017 g) in oleic acid (3.45 ml) and ODE (9.0 ml) at 300° C. The cadmium precursor solution (0.1 M) was prepared by dissolving CdO (0.1602 g) in oleic acid (3.45 ml) and ODE (9.0 ml) at 240 °C. The Zn/Cd (1/1) precursor solution (0.1 M) was prepared by dissolving ZnO (0.1017 g) and CdO (0.1602 g) in oleic acid (6.18 g) together with ODE (18.0 ml) at 300 °C. Sulfur precursor solution (0.1 M) was prepared by dissolving sulfur in ODE (9.0 ml) at 180 °C. All of the solutions were freshly prepared under N_2 atmosphere. The Cd, Zn, and Cd/Zn-precursor solutions were kept at about 90–100 °C, while the sulfur solution was kept at room temperature. For each shell growth, a calculated amount of a given precursor solution was injected with a syringe using standard air-free procedures. The amount of injection solution for each monolayer (ML) can be deduced from a calculation of the number of the surface atom [37,38].

2.4. Core-shell Q-dots synthesis using the SILAR technique

The SILAR method was used to grow multi-shells around a given CdSe-core particles, upon slightly changing the reaction conditions as reported by Li et al. [36]. To preparing CdSe/CdS/ZnS core-shell Q-dots, 3 ml of ODE, 1.23 ml of HDA were loaded into a 100 ml three-headed reaction vessel and heated up to 120 °C under vacuum for about 1 h. At this temperature, a calculated amount of CdSe–NCs (dispersed in toluene, 3.5 nm) were added and kept for about 30 min under an inert atmosphere so that toluene and other undesired material can be removed. Again, successively the solution was heated up to 240° C at which shell growth was performed. For the growth of the first monolayer (1 ML), an equal amount (0.33 ml) of each Cd and sulfur precursors were added. For the second monolayer (2 MLs) of CdS, the temperature of the Q-dots solution mixture was slightly increased (about 10° C) from 240° C. Again an equal amount of (0.33 ml) Cd and S precursors was added and kept for about 10 min to complete the growth of the 2 MLs of CdS. In the synthesis of CdSe/CdS/ZnS of 2 MLs of CdS and 1 ML of ZnS, the temperature of the reaction mixture was initially decreased to 180° C at which 0.22 ml of TOP was added and then the reaction mixture was kept for about half an hour. After that, the temperature of the solution mixture was increased up to 250° C at which an equal amount (0.45 ml) of Zn and S precursors were added. After 10 min, again 0.45 ml of Zn and S precursors were added to grow the 2MLs of ZnS. Then the temperature of the solution was kept nearly 260° C for about half an hour; then the reaction mixture was brought to the room temperature. To remove the unreacted compound and byproduct from the eluent 20 ml of methanol was added and centrifuged at 6500 rpm, this process was repeated 3–4 times and finally obtained pure CdSe/CdS/ZnS (2MLs each) Q-dots.

2.5. Preparation of the water-soluble Q-dots

The necessity to prepare water soluble Q-dots is the requirement for biological applications, as the biological environments is water. To prepare the water soluble CdSe/CdS/ZnS Q-dots, the ligand exchange or phase transformation method was used [22]. In general, the CdSe/CdS/ZnS core-shell Q-dots prepared by organic solvents are not a water-soluble because of the presence of TOP and oleic acid ligands, which are hydrophobic in nature, due to which the synthesized CdSe/CdS/ZnS Q-dots are insoluble in water. This

hydrophobic ligand could be replaced by some water soluble bi-functional molecules in which one end connect to quantum dots, and the other end is hydrophilic and may also be reactive to the molecules, as shown in Scheme 1. The TOP and oleic acid coated CdSe/CdS/ZnS (2MLs) core-shell Q-dots were dissolved in chloroform and prepared concentrated solution keeping optical density ~2 and added water solution containing MPA of equal volume (~2 ml) under vigorous stirring. The amount of MPA was adjusted roughly about 150% of the total Cd atoms on the particle surface. After 3 h, the NCs transferred into the water phase, but the solution was not optically clear at the pH range of 5–8, possibly due to interparticle hydrogen bonding between the carboxyl functions of the ligands. To remove excess MPA from the supernatant solution, centrifugation and decantation were carried out. Successive redispersion of the particles into water, especially at pH 10.5, produces a clear solution.

2.6. Experimental techniques

All the optical spectroscopic experiments were performed in toluene and water (for water soluble CdSe Q-dots) at room temperature. Absorption spectra were recorded with the UV/VIS/NIR spectrometer, Lambda-750, Perkin-Elmer. Photoluminescence (PL) spectra were collected by Horiba Jobin Yvon Fluorolog-3 spectrofluorometer equipped with double-grating at excitation as well as at emission monochromators (1200 grooves/mm) and Hamamatsu R928P photomultiplier tube (PMT). The excitation source was a 450 Watt CW xenon lamp. An instrumental correction was applied to all PL spectra to correct for the wavelength dependent PMT response and variation in output intensity from the lamp and grating efficiencies. Excitation and emission slit width were selected to be 1.0 nm and same for all measurements, and the excitation wavelength was 450 nm. Fluorescence decays were recorded using time-correlated single photon counting (TCSPC) system (DeltaFlex-01-DD, Horiba Jobin Yvon IBH Ltd) coupled with DeltaDiode Laser (450 nm) and PMT (PPD.850), and data were analyzed using the least square fitting method. High-resolution transmission electron microscopy (HRTEM) images were recorded by Tecnai G^2 20 electron microscope operated at an accelerating voltage of 200 KV. HRTEM sample was prepared on 300-mesh carbon coated Cu grids, a very dilute CdSe Q-dots nanoparticles which are dissolved in toluene is put onto carbon-coated Cu grids and after that dried the sample. The energy-dispersive X-ray images were recorded by Hitichi-640 integrated with scanning electron microscope (SEM) at accelerating voltage of 15 KV and X-Ray diffraction (XRD) pattern were recorded by BRUKER- D8.

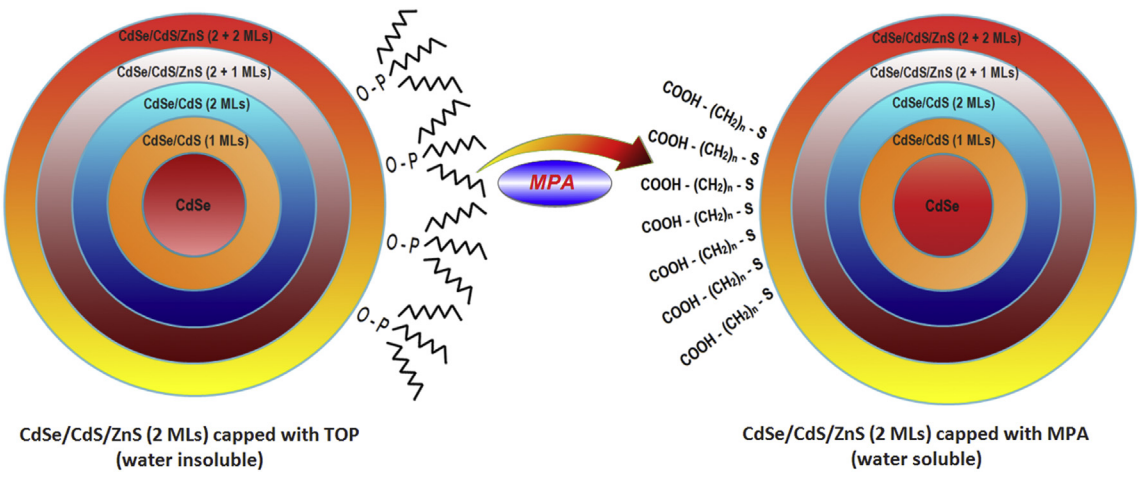
3. Results and discussion

The core-multi-shell CdSe/CdS/ZnS (2MLs of each CdS and ZnS) Q-dots were synthesized using SILAR method at an optimized temperature range of 240–260° C. The multi-shell were grown in consecutive manners over core Q-dots, and the resulting products were examined using XRD, TEM and optical spectroscopy.

Fig. 1 (a, b) shows HRTEM images of the CdSe-core and core-multi-shell Q-dots. The core-shell structure composed with two monolayers of each CdS and ZnS. With increasing the shell monolayers resulting into the increase of the diameter of the shell, which is clearly appread in size histogram (as Fig. 1a',b') of bare CdSe and CdSe/CdS/ZnS (2MLs) Q-dots. The mean diameter of bare CdSe Q-dots is about 3.4 nm and increased up to 5.3 nm after adding two monolayers of CdS and ZnS. Though the crystallinity remains same, and a shows an excellent crystallinity with continuous lattice fringes throughout the entire particle core-shell Q-dots. The blurry appearance at the rim of this particle (Fig. 1(a and b)) might be due to the molecular ligands.

The core-multi-shell structure of CdSe/CdS/ZnS (2MLs) is further examined by recording the crystallographic properties. Fig. 2 shows XRD patterns of CdSe and CdSe/CdS/ZnS (2MLs each) Q-dots. The obtained peak positions are similar to that of the bulk wurtzite lattice structure. The particle size calculated using Scherrer's formula is approximately 3.7 for CdSe and 5.7 nm for CdSe/CdS/ZnS and are quite similar to the particle size estimated from HRTEM. Further, to identify the elemental and chemical compositions of the particles, EDX spectra were measured. Table 1 and Fig. 3 show the percentage amount of each element present in the CdSe core and core-shell compositions.

Fig. 4 shows the absorption spectra of CdSe and CdSe/CdS/ZnS (2MLs each) core-multi-shell Q-dots together with the absorption spectra of intermediate products, CdSe/CdS (1 ML), CdSe/CdS (2MLs) and CdSe/CdS/ZnS (2MLs+1 ML). The absorption spectrum of core CdSe is peaking at around 538 nm. After each CdS or ZnS shell growth, the absorption spectrum shifted towards red. For CdSe/CdS/ZnS (2MLs), the absorption spectrum shifts about 20 nm and peaked at 558 nm with respect to the absorption of CdSe core. At the same time, the spectrum maintains their overall shape, indicating the narrow size distribution of the core and core-shell Q-dots, as also evident by Fig. 1(a and b). When the outer layer of CdSe



Scheme 1. Schematic representation of conversion of hydrophobic core-multi-shell (CdSe/CdS/ZnS) Q-dots to hydrophilic.

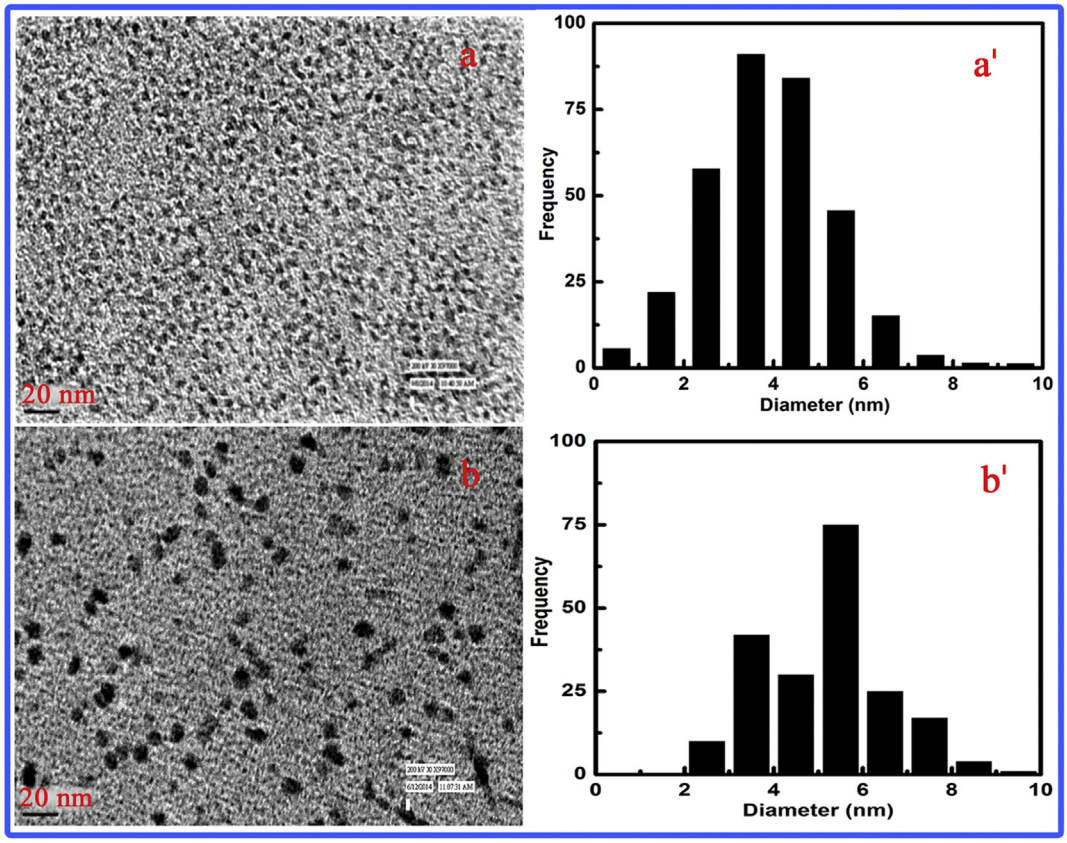


Fig. 1. The HRTEM images (a,b) and its size histogram (a',b') of CdSe and CdSe/CdS/ZnS (2MLs) core-multi-shell Q-dots.

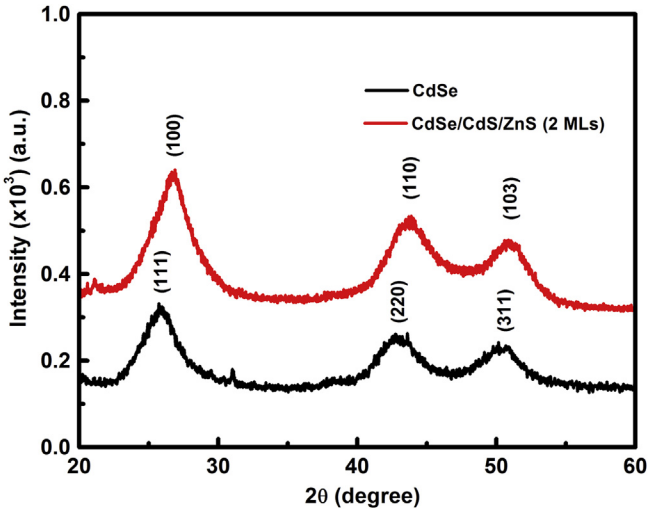


Fig. 2. The XRD patterns of CdSe-core and core-multi-shell CdSe/CdS/ZnS (2MLs) Q-dots. The peak identified with the standard JCPDS.

coated with CdS/ZnS, the red shift in absorption spectrum appeared, which may be due to the massive extension of the electronic wave function leaking out into the shell material [39].

Fig. 5 shows the PL spectra of core CdSe and core-multi-shell CdSe/CdS/ZnS (2MLs) Q-dots together with the PL spectra of intermediate products, CdSe/CdS (1 ML), CdSe/CdS (2MLs) and CdSe/CdS/ZnS (2MLs+1 ML). The excitation wavelength was 450 nm. The PL spectra are quite narrow and show an apparent red shift upon

Table 1
The percentage constituents of each element in the CdSe and CdSe/CdS/ZnS (2MLs) core-multi-shell Q-dots.

Samples	Elements	Weight %	Atom %
CdSe	Cd	66.32	58.04
	Se	33.68	41.96
CdSe/CdS/ZnS (2MLs)	Cd	56.63	43.77
	Se	25.46	24.82
	Zn	08.62	12.97
	S	07.58	16.79

increase the number of monolayers of shells. The PL maximum shifts from 559 to 581 nm. The FWHM is increased slightly with increasing the shell layers. The typical FWHM is 32.73, 32.90, 33.30, 33.80 and 34.33 nm for CdSe, CdSe/CdS (1 ML), CdSe/CdS (2MLs), CdSe/CdS/ZnS (2+1MLs) and CdSe/CdS/ZnS (2MLs each), respectively. The Stokes shift reduced from 23 nm to 17 nm upon increasing the number MLs of shell over core CdSe. The estimated QY (as described in ref. 22) of CdSe and CdSe/CdS/ZnS (2MLs) Q-dots is about 25 and 40%, respectively. However, even with TOPO the reported value of QY is about 30–50% only [10,16–18]. It has been considered that the obtained small Stoke's shift and PL FWHM indicates the band edge transition as the predominant emission [40].

Figs. 6 and 7 show the absorption and PL spectra of water-soluble CdSe/CdS/ZnS (2MLs) Q-dots at different pH with a constant CdSe/CdS/ZnS (2MLs each) concentration. PL spectra were obtained with 450 nm excitation wavelength. The size of original CdSe/CdS/ZnS Q-dots used to prepare water-soluble Q-dots was 5.3 nm and show absorption and PL peaks at 558 and 581 nm,

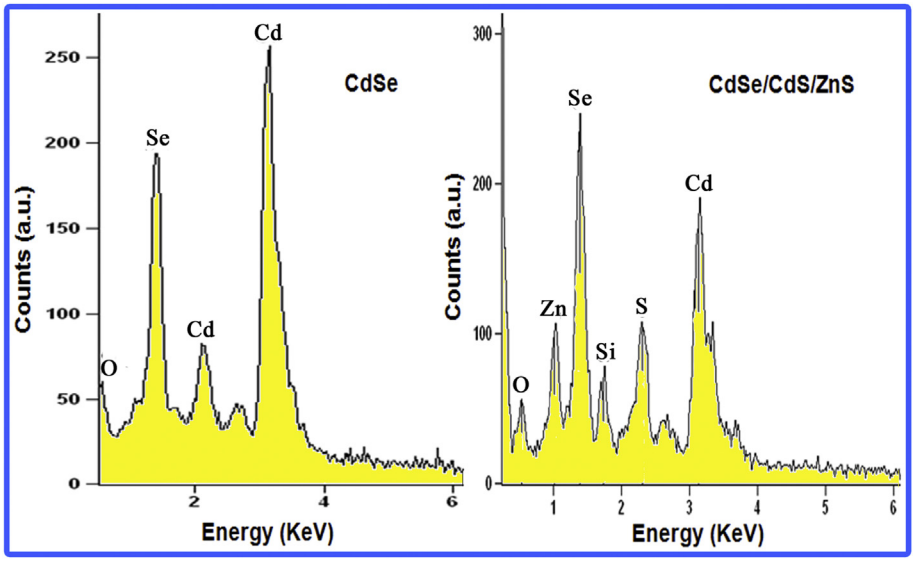


Fig. 3. The EDX spectra of CdSe and core-multi-shell CdSe/CdS/ZnS (2MLs) Q-dots.

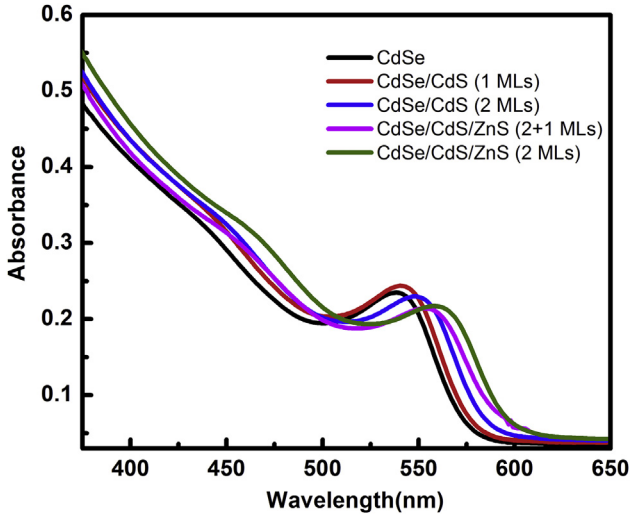


Fig. 4. The absorption spectra of CdSe and core-multi-shell Q-dots dispersed in toluene.

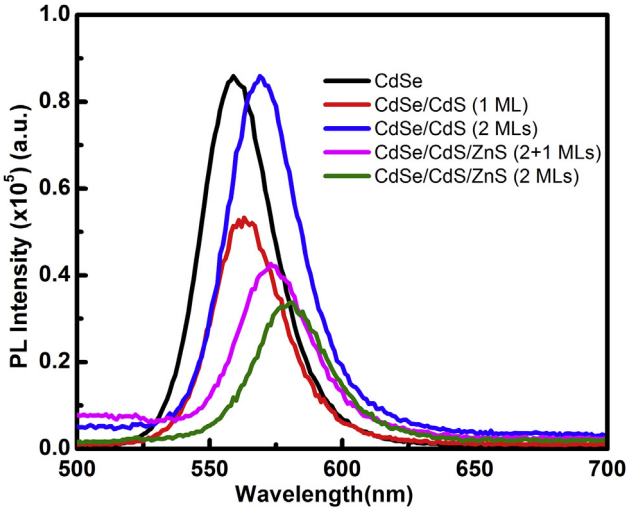


Fig. 5. The PL spectra of CdSe and core-multi-shell Q-dots dispersed in toluene at room temperature.

respectively. At different pH, the absorption maximum remains same. A similar trend was observed in the PL spectra, no significant shift in PL peak, however, the PL intensity changed with varying pH from 9.5 to 11.5. At pH 11.5, the PL intensity is maximum. At lower and higher pH, the PL spectrum does not show considerable intensity fluctuation. The estimated QY increases with increasing the pH of water (from 9.5 to 11.5), as given in Table 2. A similar trend was observed in the case of core CdSe Q-dots [22]. At basic pH, the affinity of CdSe/CdS/ZnS Q-dots is more favorable due to deprotonation of the thioalkyl acid group, which becomes negatively charged and endeavors the coordination between the ligands and the semiconductor surface giving rise to an improvement in nanoparticles dispersibility and higher QY [22]. Thus, the study reflects that the prepared water-soluble CdSe/CdS/ZnS Q-dots may be suitable for bio-application.

Fig. 8 shows the time-resolved PL decay curves of CdSe core and CdSe/CdS/ZnS (2MLs each) core-multi-shell Q-dots. The excitation wavelength was 450 nm and monitored at the peak wavelength of steady-state PL spectra. The decay curves could not fit with single exponential and are well fitted with a multi-exponential model described by the equation as follows:

$$I(t) = \sum_i A_i \exp\left(t/\tau_i\right) \quad (1)$$

where τ_i represent the time constants, and A_i represent the corresponding pre-exponential factors, respectively. Average lifetime τ_{Av} was calculated using the equation given below.

$$\tau_{av} = \sum_i A_i \tau_i / \Sigma_i A_i \quad (2)$$

The estimated lifetime and corresponding pre-exponential factor are summarized in Table 3. The decays of core and core-shell Q-dots are well fitted with four decay components, denoted by τ_1, τ_2, τ_3 and τ_3 with a pre-exponential factors of A_1, A_2, A_3 and A_4,

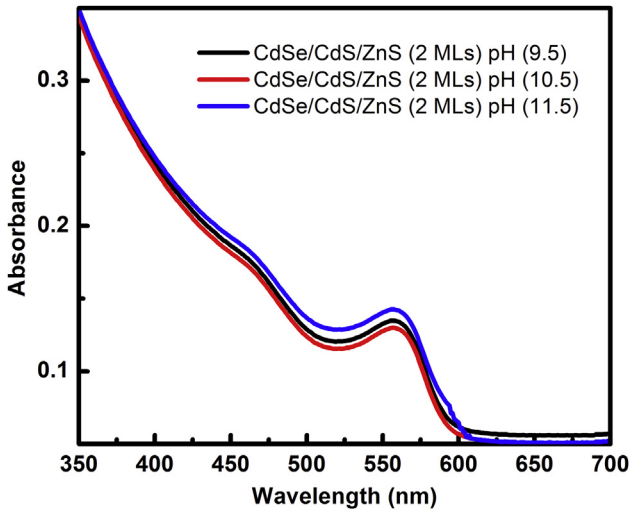


Fig. 6. The absorption spectra of water soluble core-multi-shell CdSe/CdS/ZnS (2MLs) Q-dots measured at various pH.

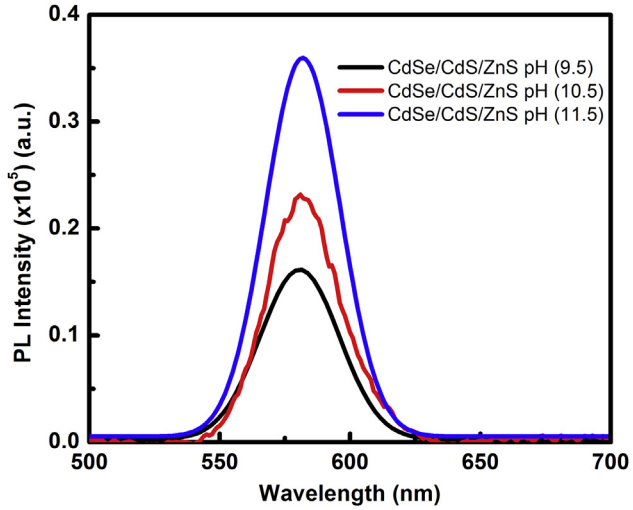


Fig. 7. The PL spectra of water soluble core-multi-shell CdSe/CdS/ZnS (2MLs) Q-dots measured at various pH.

Table 2
The quantum yield of water-soluble CdSe/CdS/ZnS (2MLs) at various pH.

S.No	Sample	Quantum yield (%)
1	CdSe/CdS/ZnS (pH) 9.5	08
2	CdSe/CdS/ZnS (pH) 10.5	10
3	CdSe/CdS/ZnS (pH) 11.5	15

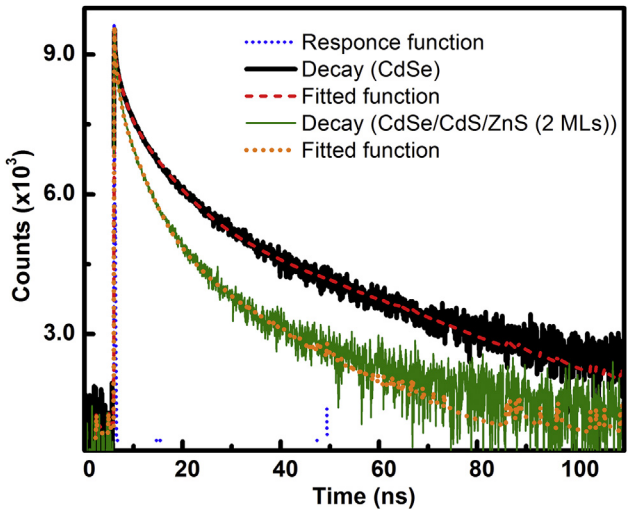


Fig. 8. Time-resolved PL decay curves of CdSe (solid black line) and core-multi-shell CdSe/CdS/ZnS (2MLs) (solid green line) Q-dots dispersed in toluene together with fitted curves (dotted line) and instrumental response function (blue dotted line). The excitation wavelength was 450 nm and monitored at peak wavelength of respective PL spectrum. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

respectively. The shortest and the longest time component are about 0.35 ns and 109.74 ns for CdSe and 0.26 ns and 86.2 ns for CdSe/CdS/ZnS (2MLs each), respectively. The average lifetime estimated from equation (2) of CdSe core and CdSe/CdS/ZnS core-shell Q-dots comes out to be 44.82 and 19.02 ns, respectively. The observed multi-exponential lifetime is correlated to the size distribution of Q-dots. However, the decreased in the average lifetime of CdSe/CdS/ZnS (2MLs each) core/shell compare to the CdSe core is due to the transition of the excited carrier to the surface quenching defects [41]. The average lifetime of core-shell is decreased dramatically and become nearly half of the CdSe core. This shows that the PL decay time significantly affected by changing surface condition of the Q-dots and upon modified with different passivation of CdS and ZnS. On the other hand, the PL QY increases nearly two folds for core-multi-shell Q-dots as compared to the core Q-dots. It has been mentioned that increase in the lifetime does not derive from the formation of type II electronic structure owing to the substantial reduction in QY [42]. The decrease in lifetime and increase in QY are consistent with the resulting extension in carrier interaction [42]. Thus, the above discussion supports a mechanism where Zn atoms serve as surface trapping defect sites, and CdSe has created less defect level because CdSe has a longer lifetime than CdSe/CdS/ZnS (2 MLs each).

4. Conclusions

The core CdSe and core-multi-shell CdSe/CdS/ZnS) Q-dots were prepared *via* hot injection method at an optimum temperature of 240° and 260° C, respectively. The prepared core and core-shell Q-

Table 3
Lifetimes and pre-exponential factors of core and core-multi-shell Q-dots dispersed in toluene.

Q-dots	(λ_{ex}) nm	(λ_{em}) nm	τ_1 (ns)	τ_2 (ns)	τ_3 (ns)	τ_4 (ns)	A_1	A_2	A_3	A_4	τ_{Av} (ns)
CdSe	450	558	0.35 ± 0.24	6.26 ± 0.30	27.60 ± 0.26	109.74 ± 0.74	3.70	17.68	51.86	26.78	44.82
CdSe/CdS/ZnS (2 MLs)	450	582	0.26 ± 0.020	4.53 ± 0.10	15.80 ± 0.21	86.20 ± 1.07	9.70	41.53	36.62	13.15	19.02

dots which are grown in an epitaxial manner exhibits narrow size distribution and controlled shape and size (3.4–5.3 nm). The QY of both core and core-multi-shell Q-dots have been improved significantly (nearly two-fold) however, the average lifetime decreased almost half and supported the given mechanism [42]. In the case of water-soluble Q-dots, the QY is maximum at higher pH.

Acknowledgements

Author thanks, Department of Science and Technology (DST) for providing time-correlated single photon counting (TCSPC) system under FIST project (SR/FST/PSI-164/2011(C)).

References

[1] M.L. Schipper, G. Iyer, A.L. Koh, Z. Cheng, Y. Ebenstein, A. Aharoni, S. Keren, L.A. Bentolila, J. Li, J. Rao, X. Chen, U. Banin, A.M. Wu, R. Sinclair, S. Weiss, S.S. Gambhir, Small 5 (2009) 126–134.
[2] J.P. Ryman-Rasmussen, J.E. Riviere, N.A. Monteiro-Riviere, Toxicol. Sci. 91 (2006) 159–165.
[3] M.S. Mehata, Appl. Phys. Lett. 100 (2012) 151908-4.
[4] M.S. Mehata, Sci. Rep. 5 (2015) 12056-11.
[5] S.S.R.K. Challa, Nanotechnologies for the Life Sciences, Wiley-VCH, Weinheim, 2005, p. 1.
[6] X. Gao, L.W.K. Chung, S. Nie, Mol. Biol. 374 (2007) 135–145.
[7] M.S. Mehata, M. Majumder, B. Mallik, N. Ohta, J. Phys. Chem. C 114 (2010) 15594–15601.
[8] U. Resch-Genger, M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke, T. Nann, Nat. Methods 5 (2008) 763–775.
[9] L. Spanhel, M. Haase, H. Weller, A. Henglein, J. Am. Chem. Soc. 109 (1987) 5649–5655.
[10] D.V. Talapin, A.L. Rogach, A. Kornowski, M. Haase, H. Weller, Nano Lett. 1 (2001) 207–211.
[11] Y.A. Wang, J.J. Li, H.Y. Chen, X.G. Peng, J. Am. Chem. Soc. 124 (2002) 2293–2298.
[12] J. Aldana, Y.A. Wang, X.G. Peng, J. Am. Chem. Soc. 123 (2001) 8844–8850.
[13] J.E.B. Katari, V.L. Colvin, A.P. Alivisatos, J. Phys. Chem. 98 (1994) 4109–4117.
[14] A.R. Kortan, R. Hull, R.L. Opila, M.G. Bawendi, M.L. Steigerwald, P.J. Carroll, L.E. Brus, J. Am. Chem. Soc. 112 (1990) 1327–1332.
[15] A. Mews, A. Eychmuller, M. Giersig, D. Schooss, H. Weller, J. Phys. Chem. 98 (1994) 934–941.
[16] X.G. Peng, M.C. Schlamp, A.V. Kadavanich, A.P. Alivisatos, J. Am. Chem. Soc. 119 (1997) 7019–7029.
[17] M.A. Hines, P. Guyot-Sionnest, J. Phys. Chem. 100 (1996) 468–471.
[18] B.O. Dabbousi, J. Rodriguez-Viejo, F.V. Mikulec, J.R. Heine, H. Mattoussi, R. Ober, K.F. Jensen, M.G. Bawendi, J. Phys. Chem. B 101 (1997) 9463–9475.
[19] D.U. Lee, D.H. Kim, D.H. Choi, S.W. Kim, H.S. Lee, K. Yoo, T.W. Kim, Opt. Express 24 (2015) A350–A357.
[20] A.O. Yalcin, B. Goris, R.J.A.V. Dijk-Moes, Z. Fan, A.K. Erdamar, F.D. Tichelaar, T.J.H. Vlugt, G.V. Tendeloo, S. Bals, D. Vanmaekelbergh, H.W. Zandbergena, M.A.V. Huise, Chem. Commun. 51 (2015) 3320–3323.
[21] D.V. Talapin, I. Mekis, S. Gotzinger, A. Kornowski, O. Benson, H. Weller, J. Phys. Chem. B 108 (2004) 18826–18831.
[22] R.K. Ratnesh, M.S. Mehata, AIP Adv. 5 (2015) 097114-10.
[23] B.K. Pong, B.L. Trout, J.Y. Lee, Langmuir 24 (2008) 5270–5276.
[24] S. Jeong, M. Achermann, J. Nanda, S. Ivanov, V.I. Klimov, J.A. Hollingsworth, J. Am. Chem. Soc. 127 (2005) 10126–10127.
[25] V. Fomenko, D.J. Nesbitt, Nano Lett. 8 (2008) 287–293.
[26] A.M. Munro, I.J.L. Plante, M.S. Ng, D.S. Ginger, J. Phys. Chem. C 111 (2007) 6220–6227.
[27] D.E. Gomez, J. van- Embden, J. Jasieniak, T.A. Smith, P. Mulvaney, Small 2 (2006) 204–208.
[28] G. Kalyuzhny, R.W. Murray, J. Phys. Chem. B 109 (2005) 7012–7021.
[29] C. Bullen, P. Mulvaney, Langmuir 22 (2006) 3007–3013.
[30] S.F. Wuister, C. deMello Donega, A. Meijerink, J. Phys. Chem. B 108 (2004) 17393–17397.
[31] C. Xu, K. Xu, H. Gu, X. Zhong, Z. Guo, R. Zheng, X. Zhang, B. Xu, J. Am. Chem. Soc. 126 (2004) 3392–3393.
[32] R. Hong, N.O. Fischer, T. Emrick, V.M. Rotello, Chem. Mater. 17 (2005) 4617–4621.
[33] D.B. Robinson, H.H.J. Persson, H. Zeng, G. Li, N. Pourmand, S. Sun, S.X. Wang, Langmuir 21 (2005) 3096–3103.
[34] J.M. Haremza, M.A. Hahn, T.D. Krauss, S. Chen, J. Calcines, Nano Lett. 2 (2002) 1253–1258.
[35] S.J. Rosenthal, I. Tomlinson, E.M. Adkins, S. Schroeter, S. Adams, L. Swafford, J. McBride, Y. Wang, L.J. Defelice, R.D. Blakely, J. Am. Chem. Soc. 124 (2002) 4586–4594.
[36] J.J. Li, Y.A. Wang, W.Z. Guo, J.C. Keay, T.D. Mishima, M.B. Johnson, X.G. Peng, J. Am. Chem. Soc. 125 (2003) 12567–12575.
[37] O. Schmelz, A. Mews, T. Basche, A. Herrmann, K. Mullen, Langmuir 17 (2001) 2861–2865.
[38] W.W. Yu, L.H. Qu, W.Z. Guo, X.G. Peng, Chem. Mater 15 (2003) 2854–2860.
[39] D. Schooss, A. Mews, A. Eychmuller, H. Weller, Phys. Rev. B 49 (1994) 17072–17078.
[40] A.M. Al-Amri, S.J. Yaghmour, W.E. Mahmoud, J. Cryst. Growth 334 (2011) 76–79.
[41] U. Kaiser, D. Jimenez de Aberasturi, R. Malinowski, F. Amin, W. J.Parak, W. Heimbrodt, Appl. Phys. Lett. 104 (2014) 041901-4.
[42] H. Zhou, G. Zhou, J. Zhou, D. Xu, X. Zhang, P. Kong, Z. Yu, RSC Adv. 4 (2014) 42316–42325.

Thermodynamic Evaluation of PTC based Organic Rankine Cycle for Power & Cooling

Devendra Kumar Gupta, Rajesh Kumar and Naveen Kumar

***Abstract*—The paper focuses on thermodynamic evaluation (first and second law) for PTC based organic Rankine cycle with single ejector (EORC) and organic Rankine cycle with double ejector (DEORC) using refrigerant R141b as working fluid and Therminol VP1 as heat transfer fluid. Thermal energy storage tanks are used to store the PTC heat, which provides the continuous net power output and cooling during insufficient solar radiation. The thermodynamic evaluation is carried out to evaluate the performance of PTC field based EORC & DEORC system which produces power and cooling simultaneously. Parametric analyses of EORC and DEORC show that inlet temperature and pressure of turbine at various extraction ratio has the significant effect on first & second law efficiency and cooling to power ratio of this system. It is also found that, annual average energy output of EORC and DEORC is 4710.57 MWh and 5104.97 MWh respectively while annual average exergy output of EORC and DEORC is 3893.86 MWh and 3436.27 MWh respectively.**

***Keywords*—Solar energy; PTC Field; Ejector; First Law Efficiency; Second Law Efficiency: Entrainment Ratio; EORC; DEORC; Cooling/Power Ratio.**

I. INTRODUCTION

As the depletion of conventional energy sources is going on and due to the use of fossil fuel for producing the power cause the emissions of green-house gases like CO_2, CO, NOx etc. which results in global warming. The conventional methods for the production of refrigeration use chlorofluorocarbons (CFCs) and hydro-fluorocarbon (HCFC) refrigerants which causes Ozone depletion and consequently green-house effect.

The renewable energy i.e. solar thermal can be used instead of fossil fuels to reduce the CO_2 emissions. Concentrated solar power (CSP) technologies (solar power tower, paraboloid dish, parabolic trough collector and linear Fresnel reflectors) are used to produce the thermal energy of high temperature and transfer it to the heat transfer fluid [1]. Concentrated solar power plants are gaining increasing interest, by using the parabolic trough collector system due to their advantages in terms of light weight in construction, lower operating cost and higher efficiency. Parabolic trough collector (PTC) technology is the best technology among all CSP technologies for power generation. Reference [2] discussed on exergy analysis and parametric study of PTC on different mass flow rates and concentration ratios for power generation. Their result shows that the exergy output, exergetic and thermal efficiencies are enhanced as the solar intensity increases. Reference [3] presented an exergy analysis of parabolic trough solar receiver. Their result shows that inlet temperature of heat transfer fluid and solar irradiance has significant effect on the thermal and exergetic performance of the PTC. As inlet temperature of heat transfer fluid increases the exergy efficiency increases but energy efficiency decreases.

Due to the variation in day-night solar irradiation, it is not possible to run the power plant continuously using CSP technology alone. The working of CSP can be enhanced by the incorporation of thermal energy storage (TES) system. Solar thermal energy can be stored during the peak sunshine hours and utilized in non-availability of sunshine hours as well as night cycle for the production of power. Reference [4] demonstrate the feasibility of two tank molten salt storage for parabolic trough solar plants. Their findings show that the use of thermal storage systems improve continuous operation of solar thermal plant and reduces the specific electricity cost. Reference [5] designed a solar thermal power plant with thermal storage for 1 MWe capacity at DNI of 0.6 kW/m^2. The plant produces about 0.8 MWh of energy output at a capacity factor of 15.6%. Reference [6] analyzed an organic Rankine cycle coupled with PTC using thermal storage has been optimized financially and energetically. Their findings show that the most suitable working fluid is cyclohexane for producing 1 MWe with PTC field.

In order to explore the solar thermal technology efficiently, the combined power and ejector cooling cycle had been proposed by many researchers. Reference [7] developed a computer programme and studied the effect of variation of performance parameters on combined power and cooling cycle. It is shown that ejector refrigeration system is more economical than absorption refrigeration system. A combined power and ejector cooling cycle was proposed by Reference [8] using R245fa as the working fluid. Their results shown that as generator temperature increases from 335 K to 415K, energy efficiency increase from 15.8% to 38.0% and exergy efficiency increase from 45.2% to 57.2% respectively. The generating temperature cannot be increased beyond a limit due to increase in turbine size. Reference [9] developed a computer program for a combined power and ejector cooling cycle using R123 as the working fluid to determine the effects of various operating parameters on the performance of the cycle. Their results show that the first and second law efficiency increase with the increase in evaporator temperature, and maximum exergy losses occurs in the boiler and ejectors. In addition to this, there is increase in first law efficiency and decrease in second law efficiency with increase in turbine inlet pressure. Reference [10] investigated the performance of a combined

Published on January 24, 2017.

D.K. Gupta, R. Kumar and N. Kumar are with Department of Mechanical Engineering, Delhi Technological University (Government of NCT of Delhi), Bawana Road, Delhi-110042, India.
(e-mail: d.k.gupta6873@gmail.com, dr.rajeshmits@gmail.com, naveenkumardce@gmail.com)

power and ejector refrigeration cycle from thermodynamic point of view and determine the optimum values of the turbine and pump inlet pressures which minimize the total thermal conductance of the cycle for the working fluid considered. Reference [11] analyses a combined power and ejector cooling system using low grade energy as heat source. The system performance were compared for five working fluid (HFE7100, HFE7000, methanol, ethanol, water) at a source temperature of 120^0C, evaporator temperature of 10^0C and condenser temperature of 35^0C. Methanol shown the highest efficiency (19.50%) followed by ethanol and water (17.30%). Their results also shown that the heat source temperature, condenser temperature and evaporator temperature have significant effect on the power output, ejector entrainment ratio and thermal efficiency of the system. An energy and exergy analyses of combined power and ejector refrigeration cycles was reported by [12]-[15] which shows that the maximum irreversibility/exergy loss occurs in heat addition process followed by the ejector and turbine.

In order to utilize the advantages of solar energy, to reduce fossil fuel consumption and alleviating environmental problems, a PTC field based EORC and DEORC is proposed for improving overall energy conversion efficiency. In EORC cooling is produce at single temperature while in DEORC cooling is produce at two different temperatures. The cycle has been analyses from both First and Second law view points and the effect of operating parameters on the performance is observed. The refrigerant R141b is used as working fluid and Therminol VP1 as heat transfer fluid. Thermal storage tanks are also used to store the thermal energy from the Sun which provides the continuous net power output during insufficient solar radiation.

II. System Description

Fig. 1a and Fig. 1b illustrates a PTC field based EORC and DEORC respectively. Fig. 2a and Fig. 2b illustrates the T-s diagram of EORC and DEORC respectively.

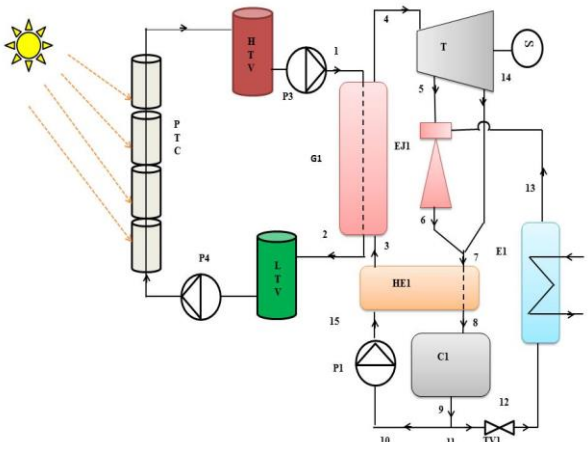


Fig. 1a Schematic diagram of Organic Rankine cycle with single ejector (EORC)

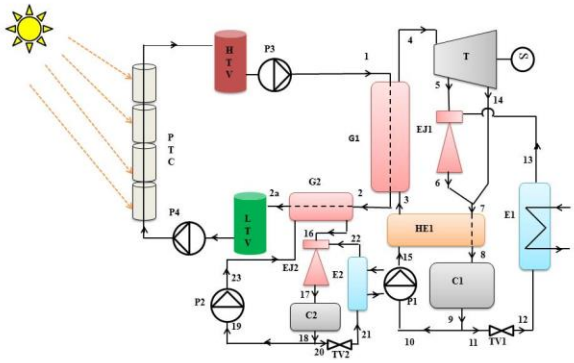


Fig. 1b Schematic diagram of Organic Rankine cycle with double ejector (DEORC)

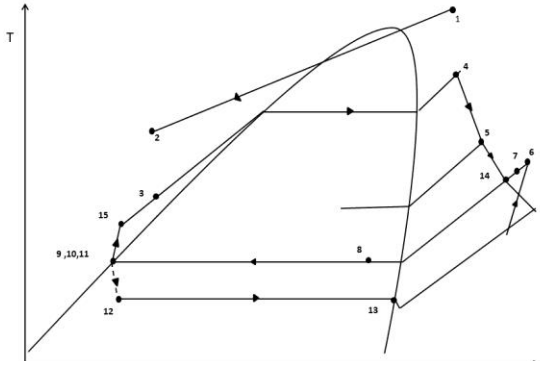


Fig. 2a T-s diagram of Organic Rankine cycle with single ejector (EORC)

The DEORC consists of parabolic trough collector (PTC) solar field, high temperature vessel, low temperature vessel, pumps, heat recovery vapor generator, turbine, ejector, condenser, throttle valve, evaporator and heat exchanger. Solar energy is used to heat the heat transfer fluid Therminol VP1 (1) with the help of PTC field. Heat transfer fluid (HTF) enters the heat recovery vapor generator (G_1) which is used to superheat the high pressure refrigerant and leaves at (2), which again enters into the heat recovery vapor generator (G_2) to vaporize the refrigerant and return back to the PTC solar field. Superheated refrigerant vapor (4) expands in the turbine. After expansion up to the extraction pressure, refrigerant vapors are extracted (5) from turbine at extraction pressure and then it flows into the ejector nozzle, entrains secondary vapor (13) from the evaporator (E_1) and mixes in mixing chamber of the ejector (EJ_1). The ejector exit stream (6) mixes with the turbine exhaust (14) and is cooled (7-8) in the heat exchanger (HE) by rejecting heat to the refrigerant leaving from the pump (P_1) which is preheated (15-3). The refrigerant leaving from heat exchanger (8) is condensed in the condenser (C1) and converts into saturated liquid (9). The saturated liquid (9) from condenser is divided into two parts. One part (11) enters into throttle valve (TV_1) and other part (10) into pump (P_1). The saturated liquid (10) pumped with the help of pump (P_1) into the heat exchanger (15). The preheated refrigerant (3) leaving from heat exchanger enters into heat recovery vapor generator (G_1) in which heat is absorbed and convert into superheated vapor (4). The saturated liquid (11) expands in the throttle valve (TV_1) to the evaporator pressure (12). The refrigerant (12) enters into the evaporator

(E_1) and produce the cooling effect by evaporation of refrigerant (12-13). The refrigerant vapour (16) flows into the ejector nozzle and entrains secondary vapor (22) from the evaporator (E_2) mixes in mixing chamber of the ejector. The ejector (EJ_2) exit stream (17) is condensed in the condenser (C_2) and converts into saturated liquid (18). Saturated liquid (18) is divided into two parts. One part (20) enters into throttle valve (TV_2) and other part (19) into pump (P_2). The saturated liquid (19) pumped with the help of pump (P_2) into the heat recovery vapor generator (G_2) and converts into saturated vapor (16). The saturated liquid (20) expands in the throttle valve (TV_2) to the evaporator pressure (21). The refrigerant (21) enters into the evaporator (E_2) and produce the cooling effect by evaporation of refrigerant (21-22). The system description of organic Rankine cycle with single ejector (EORC) is same as DEORC.

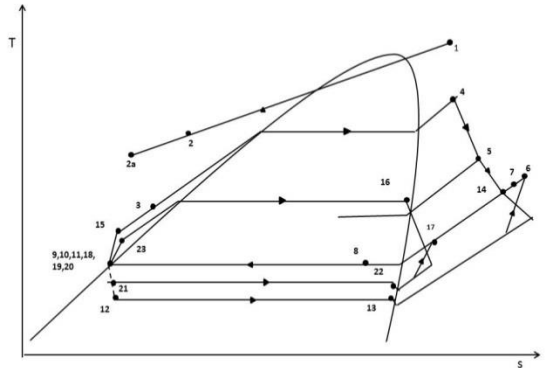


Fig. 2b T-s diagram of Organic Rankine cycle with double ejector (DEORC)

For the thermodynamic analysis, the parameters considered for the operation of EORC and DEORC are depicted in "Table I".

TABLE I: MAIN PARAMETERS CONSIDERED FOR THE ANALYSIS

Atmospheric Temperature (^{0}C)	25
Atmospheric pressure (MPa)	0.10135
Turbine inlet Temperature (^{0}C)	150
Turbine inlet pressure (MPa)	0.6
Turbine extraction pressure (MPa)	0.24
Extraction ratio	0.3
Turbine isentropic efficiency (%)	85
Pump isentropic efficiency (%)	70
Condenser temperature (^{0}C)	30
Evaporator (E_1) temperature (^{0}C)	-5
Evaporator (E_2) temperature (^{0}C)	0
Effectiveness of the heat exchanger	0.75
Aperture area (m^2)	10000
Direct normal irradiance (kWm^{-2})	0.60
HTF temperature inlet to HRVG (^{o}C)	160
HTF temperature exit to HRVG (^{o}C)	90
HRVG efficiency (%)	100
Nozzle efficiency (%)	90
Mixing chamber efficiency (%)	85
Diffuser efficiency (%)	85

III. THERMODYNAMIC ANALYSIS

Based on mass, momentum and energy equations the entrainment ratio is developed by [13].

$$\mu = \sqrt{\eta_n \eta_m \eta_d \left(h_{pf,n1} - h_{pf,2s}\right)/\left(h_{mf,ds} - h_{mf,m}\right)} - 1 \quad (1)$$

The enthalpy and entropy values for R141b at salient state points of the EORC and DEORC are taken from NIST Standard Reference Database 23, REFPROP 6.01(1998) [16].

A. First law analysis for DEORC

Solar energy received from the Sun

$$\dot{Q}_{Solar} = IA_p \quad (2)$$

where I = Direct normal irradiance (kWm^{-2}), A_p = Aperture Area (m^2)

Heat gain in the PTC field for DEORC

$$\dot{Q}_{gain} = \dot{m}(h_1 - h_{2a}) = \eta_{I,\ PTC} IA_p = \dot{m}_4(h_4 - h_3) + \dot{m}_{16}(h_{16} - h_{23}) \quad (3)$$

where $\eta_{I,\ PTC}$ = First law efficiency of PTC field

$$\eta_{I,\ PTC} = a - b\left[\frac{T_m - T_0}{I}\right] - c\left[\frac{(T_m - T_0)^2}{I}\right] \quad (4)$$

a = optical efficiency = 0.7, b = first order loss coefficient = 0.1, c = second order loss coefficient = 0, T_m = mean temperature = $(T_1+T_{2a})/2$.

The first law efficiency (η_I) of the DEORC can be defined as the ratio of the net power output ($\dot{W}_{net}$) and cooling output in both the evaporators ($\dot{Q}_{e_1}$&$\dot{Q}_{e_2}$) to the solar energy input ($\dot{Q}_{Solar}$).

$$\eta_I = \text{Energy output/Solar energy input} = \left(\dot{W}_{net} + \dot{Q}_{e_1} + \dot{Q}_{e_2}\right)/\dot{Q}_{Solar} \quad (5)$$

Turbine work:

$$\dot{W}_T = \dot{m}_4(h_4 - h_5) + \dot{m}_4(1 - E_r)(h_5 - h_{14}) \quad (6)$$

where extraction ratio: $E_r = \dot{m}_5/\dot{m}_4$ (7)

Pump work: $\dot{W}_P = \dot{m}_4(h_{15} - h_{10})$ (8)

Net power output: $\dot{W}_{net} = \dot{W}_T - \dot{W}_P$ (9)

Cooling output in evaporator (E_1)

$$\dot{Q}_{e_1} = \dot{m}_5\mu_1(h_{13} - h_{12}) \quad (10)$$

Cooling output in evaporator (E_2)

$$\dot{Q}_{e_2} = \dot{m}_{16}\mu_2(h_{22} - h_{21}) \quad (11)$$

Entrainment ratio:

Ejector (EJ_1) - $\mu_1 = \dot{m}_{13}/\dot{m}_5$ (12)

Ejector (EJ_2) - $\mu_2 = \dot{m}_{22}/\dot{m}_{16}$ (13)

B. Second law analysis for DEORC

Exergy analysis is based on second law approach. Exergy may be defined as the maximum possible reversible work obtainable in bringing the state of the system to the dead state, Mathematically

$$\dot{E} = \dot{m}[(h - h_0) - T_0(s - s_0)] \quad (14)$$

Second law efficiency of PTC field may be reported as

$$\eta_{II,\ PTC} = \{\dot{m}(h_1 - h_{2a}) - T_0(s_1 - s_{2a})\}/IA_p\psi \quad (15)$$

Maximum useful work obtainable from sun radiation (ψ) is calculated from [17] formula that is given as

$$\psi = 1 - \frac{4}{3}\frac{T_0}{T_{solar}} + \frac{1}{3}\left(\frac{T_0}{T_{solar}}\right)^4 \quad (16)$$

The second law efficiency (η_{II}) of DEORC may be reported as

$$\eta_{II} = \text{Exergy output}/\text{Exergy input} = \left(\dot{W}_{net} + \dot{E}_{e_1} + \dot{E}_{e_2}\right)/\dot{E}_{Solar} \quad (17)$$

where, $\dot{E}_{Solar}$ is exergy input associate with solar radiation falling on PTC field, $\dot{E}_{e_1}$ is the exergy associate with cooling output in the evaporator (E_1) and $\dot{E}_{e_2}$ is the exergy associate with cooling output in the evaporator (E_2),

$$\dot{E}_{e_1} = \dot{m}_{13}[(h_{12} - h_{13}) - T_0(s_{12} - s_{13})] \quad (18)$$

$$\dot{E}_{e_2} = \dot{m}_{22}[(h_{21} - h_{22}) - T_0(s_{21} - s_{22})] \quad (19)$$

$$\dot{E}_{Solar} = IA_p\psi \quad (20)$$

whereT_{Solar} is the Apparent Sun Temperature = 4500K

IV. Results and Discussion

The first and second law analyses is conducted for EORC and DEORC, their results are compared by varying operating parameters such as turbine inlet temperature (TIT) and turbine inlet pressure (TIP) at different extraction ratio on the performance of the cycle.

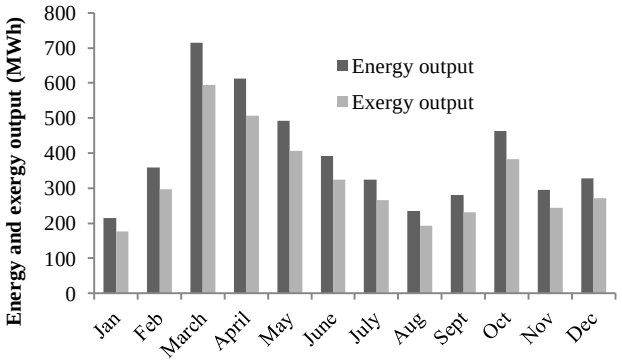


Fig. 3 Monthly energy and exergy output of EORC

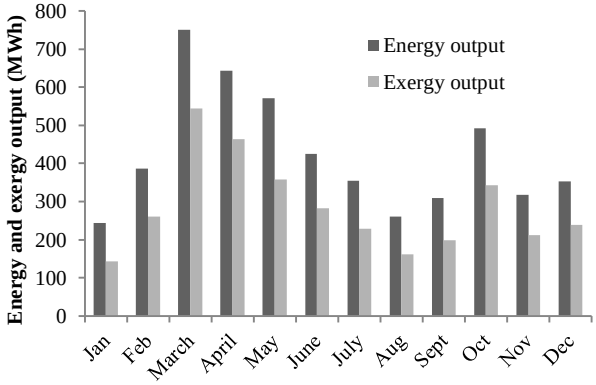


Fig. 4 Monthly energy and exergy output of DEORC

Fig. 3 and Fig. 4 show the monthly energy and exergy output of EORC and DEORC. It is observed that due to higher DNI in the month of March, April, May and October the energy and exergy output is higher while in the remaining months the energy and exergy output obtained is lower due to low value of DNI. The highest energy output (715.53 MWh) and exergy output (593.8 MWh) obtained in the month of March while lowest energy output (214.74 MWh) and exergy output (176.57 MWh) in the month of January in EORC. The maximum energy output in the month of March is 749.95 MWh and maximum exergy output 544.38 MWh while least energy output 244.07 MWh and exergy output is 142.82 MWh registered in the month of January in DEORC. The annual average energy output of EORC and DEORC is 4710.57 MWh and 5104.97 MWh respectively while annual average exergy output of EORC and DEORC is 3893.86 MWh and 3436.27 MWh respectively. It is concluded from Fig. 3 and Fig. 4 that yearly average energy output increases while average exergy output decreases with the addition of ejector refrigeration system in the EORC.

Fig. 5 shows the energy input/output and exergy input/output in EORC and DEORC at input thermodynamic parameters as shown in Table I. In EORC the total useful energy output of the cycle ($\dot{W}_{net} + \dot{Q}_{e_1}$) is 686 kW and the exergy output ($\dot{W}_{net} + \dot{E}_{e_1}$) is 571.22 kW is observed. In DEORC the total energy output ($\dot{W}_{net} + \dot{Q}_{e_1} + \dot{Q}_{e_2}$) is 711.2 kW and the exergy output ($\dot{W}_{net} + \dot{E}_{e_1} + \dot{E}_{e_2}$) is 535.26 is observed. It is clear that energy output in DEORC is higher than EORC.

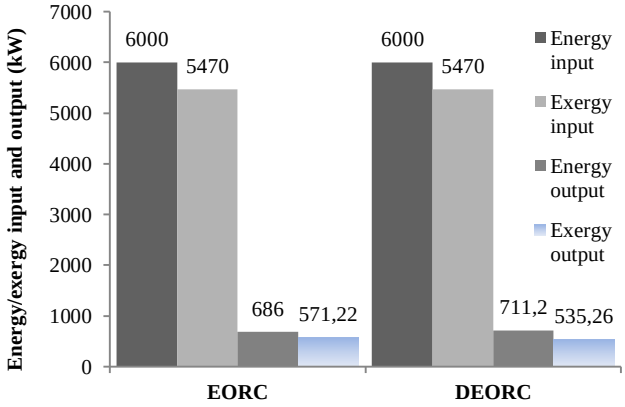


Fig. 5. Energy input/output, Exergy input/output of EORC and DEORC

Fig. 6, 7, and 8 shows the effect of turbine inlet temperature (TIT) on first law efficiency, second law

efficiency and cooling/power ratio at different extraction ratio in EORC and DEORC. As the turbine inlet temperature increases the net power output increases due to increase in enthalpy drop in the turbine and cooling output increases due to increase in the enthalpy of R141b at the inlet of the ejector. This causes the increase in velocity of stream at the exit of the nozzle ejector, results in increase in entrainment ratio hence cooling output in EORC and DEORC increases. The rate of increase of net power output is higher than rate of increase of cooling output, the cooling to power ratio decreases. Due to increase in power output and cooling output the first law efficiency increases with increase in turbine inlet temperature. The second law efficiency increases with increase in turbine inlet temperature due to increase in total exergy output. The result shows that the first law efficiency of DEORC is greater than first law efficiency of EORC. With the increase in extraction ratio at the same turbine inlet temperature, the net power output decreases and cooling output increases. Since the rate of increase in cooling output is higher than the rate of decrease in net power output. The cooling to power ratio and first law efficiency increases while second law efficiency decreases as the extraction ratio increases. Fig. 6, 7, and 8 indicate that first law efficiency and cooling to power ratio is higher and second law of efficiency is lower for DEORC than EORC at the same working conditions.

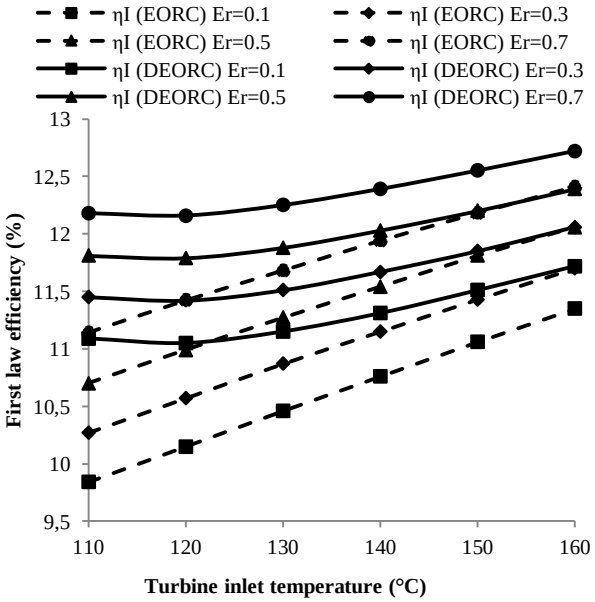


Fig. 6. Variation of first law efficiency with TIT at different extraction ratio

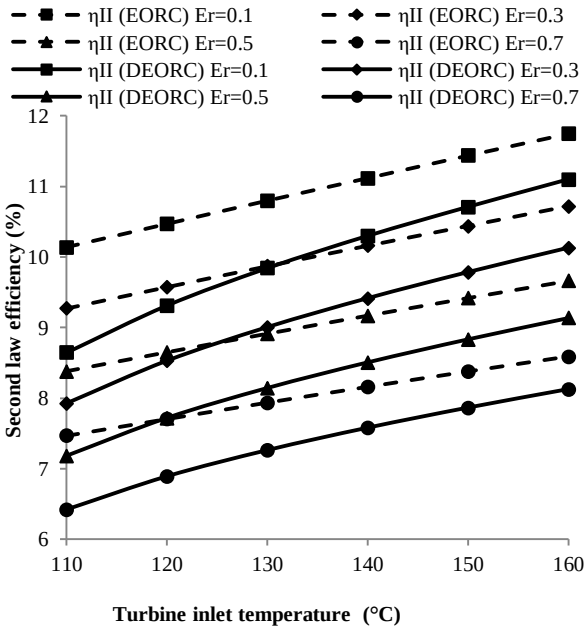


Fig. 7. Variation of second law efficiency with TIT at different extraction ratio

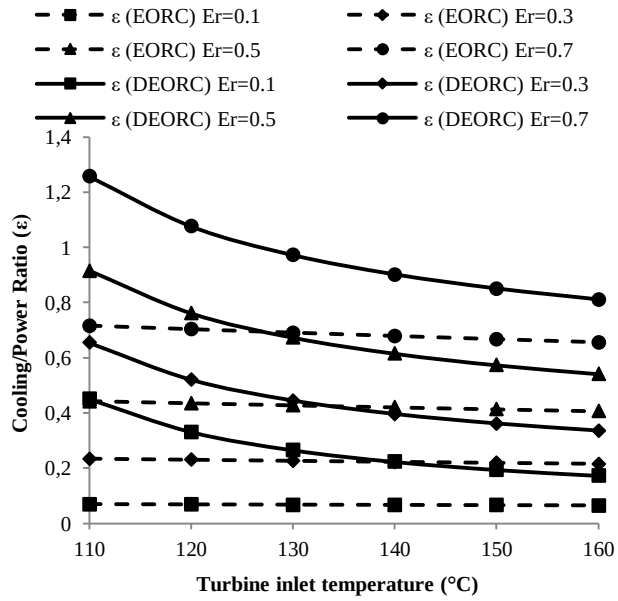


Fig. 8. Variation of Cooling/Power ratio with TIT at different extraction ratio

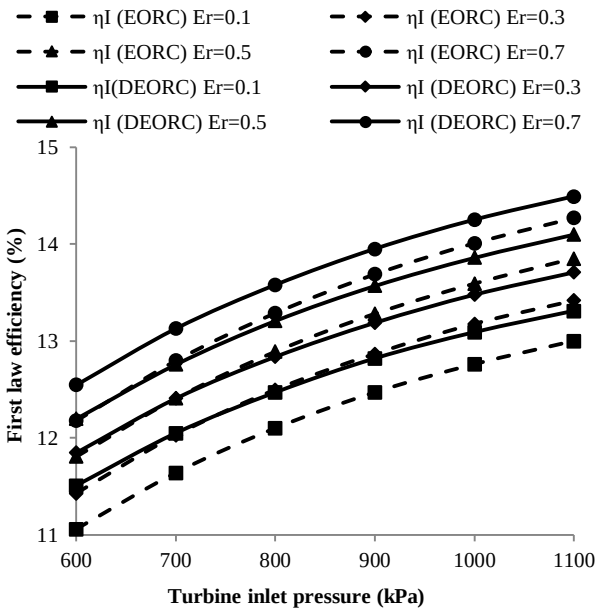


Fig. 9. Variation of First law efficiency with TIP at different extration ratio

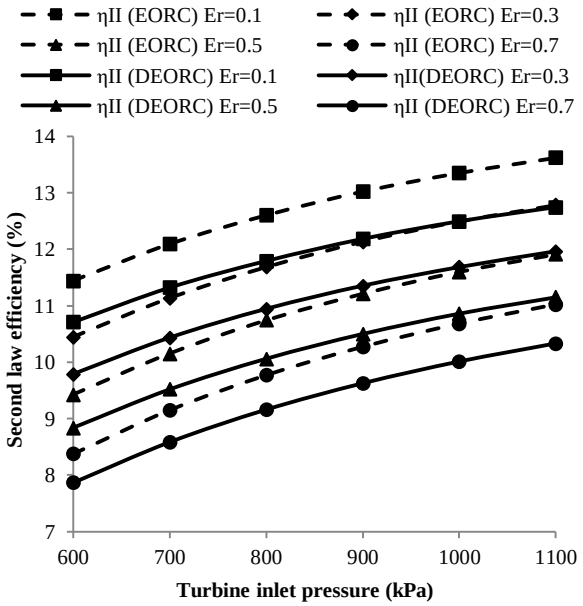


Fig. 10. Variation of Second law efficiency with TIP at different extration ratio

Fig. 9, 10, and 11 show the effect of variation of turbine inlet pressure (TIP) on first law efficiency, second law efficiency and cooling to power ratio at different extraction ratio. As the turbine inlet pressure increases, the net power output increases and cooling output decreases for EORC and DEORC. The rate of increase in net power output is greater than rate of decrease in cooling output results in increase in first law efficiency and decrease in cooling to power ratio for EORC and DEORC with increase in turbine inlet pressure. The same trend for exergy efficiency observed with increase in turbine inlet pressure as shown in Fig. 10. With the increase in extraction ratio at the same turbine inlet pressure, the net power output decreases and cooling output increases. Since the rate of increase in cooling output is higher than the rate of decrease in net power output. The cooling to power ratio and first law efficiency increases while second law efficiency decreases as the extraction ratio increases. Fig. 9, 10, and 11 indicate that first law efficiency and cooling to power ratio is higher and second law efficiency is lower for DEORC than EORC at the same working conditions.

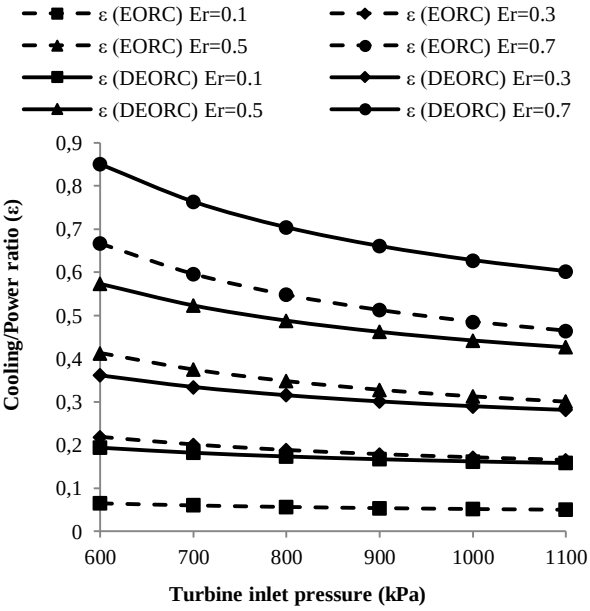


Fig. 11. Variation of Cooling/Power ratio with TIP at different extration ratio

Table II shows the performance of the EORC and DEORC. It is observed that the cooling and power output in DEORC is 711.23 kW as compared to 686 kW of EORC. The total energy output in DEORC is 11.85% of solar energy input (100%) is observed.

TABLE II: THE PERFORMANCE OF THE EORC AND DEORC.

	Organic Rankine cycle with single ejector (EORC)		**Organic Rankine cycle with double ejector (DEORC)**	
	Amount	**% of solar energy input**	**Amount**	**% of solar energy input**
Solar energy input (kW)	6000.00	100.00	6000.00	100.00
Turbine work (kW)	572.50	9.54	531.90	8.85
Total pump work (kW)	9.55	0.159	9.37	0.15
Total cooling output (kW)	123.10	2.05	188.70	3.15

Net power output (kW)	562.90	9.38	522.53	8.70
Energy output (kW)	686.00	11.43	711.23	11.85
Energy loss (kW)	5314.00	88.57	5288.77	88.15

Table III shows the distribution of exergy in EORC and DEORC. The total exergy output in DEORC is 535.26 kW as compared to exergy output of 571.22 kW in EORC.

TABLE III: THE DISTRIBUTION OF EXERGY IN EORC AND DEORC.

	EORC		DEORC	
	Amount	% of exergy	Amount	% of exergy
Exergy input (kW)	5470.00	100.00	5470.00	100.00
Exergy output (kW)	571.22	10.44	535.26	9.785
Exergy destruction/ losses (kW):				
(i) Generator	218.50	3.99	212.84	3.89
(ii) Turbine	81.46	1.49	75.68	1.38
(iii) Ejector	87.23	1.59	110.19	2.02
(iv) Condenser	60.44	1.10	63.75	1.165
(v) Pump	9.36	0.17	9.19	0.168
(vi) Heat exchanger	10.77	0.20	10.00	0.183
(vii) Throttle valve	1.65	0.03	2.24	0.041
(viii) Evaporator	4.85	0.09	5.92	0.108
(ix) Parabolic trough collector	4424.52	80.90	4444.93	81.26
Total Exergy destruction/loss (kW)	4898.78	89.56	4934.74	90.215

V. CONCLUSION

This paper presents thermodynamic analysis of PTC based EORC & DEORC using R141b as working fluid and Therminol VP1 as heat transfer fluid. The following main conclusions in this study are as follows:

- With the increase in inlet temperature of turbine (110^0C - 160^0C), the first and second law efficiency of EORC and DEORC increases while cooling /power ratio decreases.
- With the increase in the inlet pressure of turbine (600 kPa - 1100 kPa), the first and second law efficiency of EORC and DEORC increases whereas cooling /power ratio decreases.
- The first law efficiency and cooling to power ratio is higher in DEORC as compared to EORC at the same working condition.
- In EORC the maximum energy output (715.53 MWh) and the maximum exergy output (593.8 MWh) obtained in the month of March while minimum energy output (214.74 MWh) and minimum exergy output (176.57 MWh) in the month of January. The annual average energy output of EORC is 4710.57 MWh while annual average exergy output of EORC 3436.27 MWh.
- In DEORC the highest energy output (749.95 MWh) and exergy output (544.38 MWh) is registered in the month of March while least energy output (244.07 MWh) and exergy output (142.82 MWh) registered in the month of January respectively. The annual average energy output of DEORC is 5104.97 MWh while annual average exergy output of DEORC is 3436.27 MWh respectively.
- With the addition of ejector in EORC the first law efficiency increases from 11.43% to 11.85% while second law efficiency decreases from 10.44% to 9.785% at input thermodynamic parameters as shown in Table I.

APPENDIX

Nomenclature

A_p	Aperture area (m^2)
I	Direct normal irradiance (kWm^{-2})
$\dot{E}$	Exergy rate (kW)
$\dot{Q}$	Heat transfer rate (kW)
T	Absolute temperature (K), Turbine
P	Pressure (MPa), Pump
$\dot{W}$	Power output/input (kW)
h	Specific enthalpy ($kJ\ kg^{-1}$)
$\dot{m}$	Mass flow rate ($kg\ s^{-1}$)
s	Specific entropy ($kJ\ kg^{-1}K^{-1}$)
G	Heat recovery vapor generator (HRVG)
EJ	Ejector
C	Condenser
HE	Heat exchanger
TV	Throttle valve
E	Evaporator
Er	Extraction ratio
HTF	Heat transfer fluid (Therminol VP1)
EORC	Organic Rankine cycle with single ejector
DEORC	Organic Rankine cycle with double ejector
PTC	Parabolic trough collector

Greek symbols

μ	Entrainment ratio
ε	Cooling to power ratio
η	Efficiency (%)

Subscript

$\mathbf{e}_1$	Evaporator 1
$\mathbf{e}_2$	Evaporator 2
P	Pump
T	Turbine

1, 2, 3…….a, b, c…..state points in Figures 1a, 1b, 2a, 2b.

REFERENCES

[1] M. T. Dunham, and B. D. Iverson, "High-efficiency thermodynamic power cycles for concentrated solar power systems," *Renewable and Sustainable Energy Reviews*, Vol. 30, pp. 758-770, 2014.

[2] S. K. Tyagi, S. Wang, M. K. Singhal, S. C. Kaushik, and S.R. Park, "Exergy analysis and parametric study of concentrating type solar collectors," *International Journal of Thermal Sciences*, Vol. 46, pp. 1304-1310, 2007.

[3] R. V. Padilla, A. Fontalvo, G. Demirkaya, A. Mrtinez, and A. G. Quiroga, "Exergy analysis of parabolic trough solar receiver," *Applied Thermal Engineering*, Vol. 67**,** pp. 579-586, 2014.

[4] U. Herrmann, B. Kelly, and H. Price, "Two tank molten salt storage for parabolic trough solar power plants," *Energy*, Vol. 29, pp. 883-893, 2004.

[5] N. B. Desai, S. Bandyopadhyay J. K. Nayak, R. Banerjee, and S. B. Kedare, "Simulation of 1MWe Solar Thermal Power Plant," *ISES Solar World Congress*, 2013.

[6] C. Tzivanidis, E. Bellos, and K. A. Antonopoulos, "Energetic and financial investigation of a stand-alone solar-thermal organic Rankine cycle power plant," *Energy Conversion and Management*, Vol. 126, pp. 421-433, 2016.

[7] G. K. Alexis, "Performance parameters for the design of a combined refrigeration and electrical power cogeneration system," *International Journal of Refrigeration*, Vol. 30, pp. 1097–1103, 2007.

[8] B. Zheng, and Y. W. Weng, "A combined power and ejector refrigeration cycle for low temperature heat sources," *Solar Energy*, Vol. 84, pp. 784-791, 2010.

[9] M. M. Rashidi, O. A. Beg, and A. Aghagoli, "Utilization of waste heat in combined power and ejector refrigeration for a solar energy source," *International Journal of Appl. Math and Mech*., Vol. 8(17), pp. 1-16, 2012.

[10] A. Habibzadeh, M. M. Rashidi, and N. Galanis, "Analysis of a combined power and ejector-refrigeration cycle using low temperature heat," *Energy Conversion and Management*, Vol. 65, pp. 381-391, 2013.
[11] X. Chen, Y. Su, S. Omer, and S. Riffat, "Theoretical investigations on combined power and ejector cooling system powered by low-grade energy source," *International Journal of Low-Carbon Technologies Advance Access*, Vol. 0, pp. 01-10, 2015.
[12] J. Wang, Y. Dai, and L. Gao, "Parametric analysis and optimization for a combined power and refrigeration cycle," *Applied Energy*, Vol.85, pp. 1071-1085, 2008.
[13] Y. Dai, J. Wang, and L. Gao, "Exergy analysis, parametric analysis and optimization for a novel combined power and ejector refrigeration cycle," *Applied Thermal Engineering*, Vol. 29, pp. 1983-1990, 2009.
[14] D. K. Gupta, R. Kumar, and N. Kumar, "First and second law analysis of solar operated combined Rankine and ejector refrigeration cycle," *Applied Solar Energy*, Vol. 50(2), pp. 113-121, 2014.
[15] B. K. Agrawal, and M. N. Karimi, "Parametric, exergy and energy analysis of low grade energy and ejector refrigeration cycle," *International Journal of Sustainable Building Technology and Urban Development*, Vol. 4(2), pp. 170-176, 2013.
[16] NIST Standard Reference Database 23, 1998, NIST Thermodynamic and Transport Properties of Refrigerants and Refrigerants Mixtures REFPROP, Version 6.01.
[17] R. Petela, "Exergy of undiluted thermal radiation", Solar Energy, Vol.74 (6), pp. 469-488, 2003.



Devendra Kumar Gupta did his BE in Mechanical Engineering in 1991 and post-graduation, i.e. ME in Materials Handling in 2005 from Madhav Institute of Technology & Science, Gwalior, M.P., India. Besides his industrial experience of four years, he has teaching experience of 16 years in the field of Thermal Engineering. Presently, he is pursuing his PhD from Department of Mechanical Engineering, Delhi Technological University (Government of NCT of Delhi), Bawana Road, Delhi.



Dr. Rajesh Kumar did his B. Tech. in Mechanical Engineering from HBTI Kanpur in the year 1998 and ME in Thermal Engineering in the year 2000 from IIT, Roorkee. He did his PhD in the year 2008 from Jamia Millia Islamia, Delhi. He has teaching experience of more than 16 years in the field of Thermodynamics and refrigeration. Presently, he is working as an Associate Professor in the Department of Mechanical Engineering, Delhi Technological University (Government of NCT of Delhi), Bawana Road, Delhi. His several papers have been published in international journals having high impact factor. He is guiding 4 PhD scholar in Mechanical Engineering department and supervised 12 M. Tech. dissertations in the area of thermal engineering.



Dr. Naveen Kumar has more than 25 years of experience of industry, teaching and research, which includes more than 20 years at Delhi Technological University (Formerly Delhi College of Engineering), where he is currently a Professor of Mechanical Engineering. Besides teaching at UG and PG levels, Prof. Naveen Kumar has made a significant contribution to R&D in the area of alternative fuels for their utilization in Internal Combustion Engines. Apart from supervising Ph.D. and PG scholars, he has undertaken a number of major R&D Projects sponsored by the Government and Industrial Houses. He has traveled extensively to abroad for delivering invited lectures and carrying out collaborative research. His work on biofuels is well recognized in India and abroad.